Optogenetics and Computer Vision for C. elegans Neuroscience and Other Biophysical Applications

Aravinthan D. T. Samuel Andrew Michael Leifer

Optogenetics and computer vision for *C. elegans* neuroscience and other biophysical applications

Abstract

This work presents optogenetics and real-time computer vision techniques to non-invasively manipulate and monitor neural activity with high spatiotemporal resolution in awake behaving *Caenorhabditis elegans*. These methods were employed to dissect the nematode's mechanosensory and motor circuits and to elucidate the neural control of wave propagation during forward locomotion. Additionally, similar computer vision methods were used to automatically detect and decode fluorescing DNA origami nanobarcodes, a new class of fluorescent reporter constructs.

An optogenetic instrument capable of real-time light delivery with high spatiotemporal resolution to specified targets in freely moving *C. elegans*, the first such instrument of its kind, was developed. The instrument was used to probe the nematode's mechanosensory circuit, demonstrating that stimulation of a single mechanosensory neuron suffices to induce reversals. The instrument was also used to probe the motor circuit, demonstrating that inhibition of regions of cholinergic motor neurons blocks undulatory wave propagation and that muscle contractions can persist even without inputs from the motor neurons.

The motor circuit was further probed using optogenetics and microfluidic techniques. Undulatory wave propagation during forward locomotion was observed to depend on stretch-sensitive signaling mediated by cholinergic motor neurons. Specifically, posterior body segments are compelled, through stretch-sensitive feedback, to bend in the same direction as anterior segments. This is the first explicit demonstration of such feedback and serves as a foundation for under-

standing motor circuits in other organisms.

A real-time tracking system was developed to record intracellular calcium transients in single neurons while simultaneously monitoring macroscopic behavior of freely moving *C. elegans*. This was used to study the worm's stereotyped reversal behavior, the omega turn. Calcium transients corresponding to temporal features of the omega turn were observed in interneurons AVA and AVB.

Optics and computer vision techniques similar to those developed for the *C. elegans* experiments were also used to detect DNA origami nanorod barcodes. An optimal Bayesian multiple hypothesis test was deployed to unambiguously classify each barcode as a member of one of 216 distinct barcode species.

Overall, this set of experiments demonstrates the powerful role that optogenetics and computer vision can play in behavioral neuroscience and quantitative biophysics.

Contents

Dedicated to Ethan Townsend (1985-2009).

Author List

The author for Chapter 1 is Andrew M Leifer.

The authors for Chapter 2 are Andrew M Leifer, Christopher Fang-Yen, Marc Gershow, Mark J Alkema, and Aravinthan D T Samuel.

The authors for Chapter 3 are Quan Wen, Elizabeth Hulme, Sway Chen, Xinyu Liu, Marc Gershow, Andrew M Leifer, Victoria Butler, Christopher Fang-Yen, William R Schafer, George Whitesides, Matthieu Wyart, Dmitri B Chklovskii, and Aravinthan D T Samuel.

The authors for Chapter 4 are Chenxiang Lin, Ralf Jungmann Andrew M Leifer, Chao Li, Daniel Levner, William M Shih, and Peng Yin.

The authors for Chapter 5 are Andrew M Leifer, Mark C Leifer and Chenxiang Lin.

The authors for Chapter 6 are Andrew M Leifer, Christopher Clark, Mark Alkema and Aravinthan D T Samuel.

The author for Chapter 7 is Andrew M Leifer.

Detailed information about author contributions are provided at the end of each chapter in the section entitled “Manuscript Information.”

Acknowledgments

This thesis would not have been possible without the support of many individuals throughout my graduate studies. Thanks goes first and foremost to my advisor, Professor Aravinthan Samuel, who graciously took me into his lab and supported me with deft guidance and thoughtful mentorship. Aravi conducts science with an integrity and earnestness that I very much admire and it has been a privilege to learn from his example.

The Samuel lab is inherently collaborative and the work here has benefited from many others within the lab. I thank Christopher Fang-Yen for introducing me to optogenetics and for getting me started on such an exciting project. Its been a privilege to work with Quan Wen who has an incredible knack for insightful experiments. I've learned much from his lead. Marc Gershow never fails to enliven the lab and I am indebted to him for many great brainstorming sessions and for teaching me how to manage very large software projects. I have enjoyed working with Mason Klein on microscopy and optics projects. I also thank Linjiao Luo for showing me the ropes and thank Elizabeth Kane for critical graduate student solidarity.

I worked with the following undergraduates: Anji Tang, Laura Freeman and Konlin Shen. I thank them for their excellent work, patience and dedication. Sway Chen in particular helped conduct many of the experiments in Chapter 3.

This work was done collaboratively with others at Harvard and beyond. Specific contributions to each chapter are listed in the sections marked "Manuscript Information." In particular, thanks goes to Mark Alkema and Christopher Clark of the Alkema Lab at UMass Worcester. Mark provided guidance and mentorship and helped design the locomotion experiments. Chris Clark and I worked closely together on the study of the omega turns. He made many of the strains used here and I greatly appreciate his companionship during many hours of data collection.

The Samuel lab benefits from its close proximity to the Zhang lab upstairs. I have learned greatly from Yun Zhang, Adam Bahrami and Michael Hendriks.

The DNA origami work was conducted in collaboration with William Shih and spearheaded by Chenxiang Lin of the Shih lab. It has been a pleasure to work with Chenxiang, and I thank William in particular for allowing me to rotate in his lab.

I also had the unique opportunity to collaborate with my father, Mark Leifer, on the mathematical theory behind the DNA barcode detection software, which was an especially fun and rewarding experience.

Thanks to the following members of the worm community for sharing advice and reagents: Massimo Hilliard, Brent Neumann, Mei Zhen and Niels Ringstad. Thanks to Ed Boyden and Brian Chow of the Boyden lab who took the time to meet with me early on to discuss optogenetics. Thanks to Jeffrey Stirman for open, friendly and collegial competition.

I arrived at *C. elegans* neuorscience research through a circuitous route. I owe thanks to Marcus Meister who first sparked my interest in neuroscience during a seminar lecture and then allowed me to rotate in his lab. L Mahadavan and William Ryu first introduced me to the existence of *C. elegans* during a final project in Maha's class. Through them, I eventually found my way to the Samuel lab.

Prior to Harvard, certain individuals inspired me to go to graduate school. Tom Perkins and Ashley Carter of JILA at NIST CU-Boulder gave me my first foray into academic research, and both have been invaluable resources ever since. At critical junctures the following provided crucial encouragement when I needed it: Sidney Drell, Mark Kasevich, Benn Tannenbaum, Rick Pam, David Goldhaber-Gordon, Sandy Alexander, Zev Bryant and Naveen Sinha.

To Michele Jakoulov and Jim Hogle, thank you for making the Biophysics program such a supportive home. Thanks to my friends and members of my cohort: Dan Chonde, Nate Derr, Ashley Gibbs, Bryan Harada, Alison Hill, Xavier Rios, Peter Stark and Kevin Takasaki. In particular, thanks go to fellow biophysics students Benjamin Schwartz and Alex Fields for being gifted teachers and steadfast friends.

Thanks to my dissertation Advisory Committee: Adam Cohen, Ed Boyden, Mark Alkema and Markus Meister; and to my Thesis Committee: Markus Meister, Yun Zhang and Florian Engert.

This thesis is dedicted to the brilliant scientist and engineer Ethan Townsend. No person has done more to teach me to appreciate the friendship of others.

I am fortunate to have the unwavering support of my loving parents Anne and Mark Leifer. I thank them for all of the time they have spent counseling and encouraging me and for instilling in me a deep-seated value in academics and education at an early age. I also thank my brother Daniel Leifer for being a close friend.

I am a product of my family and I love them very much.

Finally, I want to thank Franziska Graf. She is a brilliant scientist, caring friend and dedicated partner. I love her dearly and this entire process has been much better with her by my side.

I received financial support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE-0644491.

Some nematode strains used in this work were provided by the Caenorhabditis Genetics Center, which is funded by the NIH National Center for Research Resources (NCRR).

Thanks to Mark Leifer and Camille Rickets for copyediting. Text is typeset in XeTeX using a modified version of Jordan Suchow's template, `http://github.com/aleifer/LaTeX-template-for-Harvard-dissertation`.

The source code for the dissertation is available at `http://github.com/aleifer/dissertation`.

To cross the threshold from where we are to where we want to be, major conceptual shifts must take place in how we study the brain. One such shift will be from studying elementary processes—single proteins, single genes, and single cells—to studying systems properties—mechanisms made up of many proteins, complex systems of nerve cells, the functioning of whole organisms, and the interaction of groups of organisms. Cellular and molecular approaches will certainly continue to yield important information in the future, but they cannot by themselves unravel the secrets of internal representations of neural circuits or the interaction of circuits—the key steps linking cellular and molecular neuroscience to cognitive neuroscience.

Eric R. Kandel, [80]

1 Introduction

How do a collection of neurons work together to receive information from the environment, encode that information, and then process it to generate purposeful behavior? This is the fundamental question confronting researchers in the field of systems neuroscience, and is the singular focus of many laboratories around the world. While much progress is being made to answer this question, our understanding has been limited in part by the lack of tools to simultaneously probe and observe the activity of individual neurons and their effect on whole-organism behavior. In particular, there is a distinct need for non-invasive techniques to stimulate and record from ensembles of neurons across an organism with single-cell resolution in awake, intact, unrestrained animals whose behavior can be observed simultaneously.

This need has motivated my thesis work and prompted me to develop new tools and techniques combining optogenetics, microscopy, and computer vision, to optically monitor and manipulate neural activity in freely behaving *Caenorhabditis elegans*. Using the tools presented in this work, I provide new insights into how a network of neurons drives locomotion. That work is covered in Chapters 2, 3 and

6. Along the way, I employed similar techniques toward the development of novel fluorescent reporter constructs for next-generation microarray and lab-on-a-chip technologies. That work is discussed in Chapters 4 and 5. The neuronal basis of behavior, however, remains the primary focus of the thesis and of this introduction.

1.1 Background

The neuron is the fundamental unit of the brain. Neurons are responsible for encoding information from an organism's environment and performing computations to transform that information into behavior. How this chain of events takes place has been studied with different approaches at different length scales, depending on the tools available at the time [50, 81].

At the end of the 19th century, the development of the Golgi stain enabled Ramón y Cajal to approach this question by studying neuron morphology. Cajal examined how individual neurons connected to one another, providing the first insights into the structure and function of neural circuits. Around the same time, the study of macroscopic lesions in the brain provided another class of insights. Researchers like Pierre Paul Broca and Carl Wernicke were able to correlate lesions in human brains with defects in cognition, and thus were able to posit macroscopic models for how different brain regions interact with the nervous system. Single neuron function was again the focus of study as a new tool– electricity –was harnessed. Luigi Galvani and Alassendro Volta in the 18th century, Hermann von Helmholtz and David Ferrier in the 19th century, and Charles Sherrington in the early 20th, used electrodes to probe the electrical activity of individual neurons. Once op-amps were developed from radar work during World War II, it became possible to clamp an individual neuron at a precise voltage or current. Alan Hodgkins, Andrew Huxley, Bernard Katz and others used voltage clamps to rigorously characterize signaling within a single neuron. In a combination of microscopic and macroscopic investigation, the recent development of laser killing and genetic engineering have allowed researchers to systematically kill and ablate neurons or classes of neurons from a neural circuit while observing the resulting behavioral defects.

None of these techniques however, have enabled researchers to non-invasively

perform cell-specific intracellular manipulations across an entire organism while simultaneously observing unrestrained behavior. Such an approach is ideal to address the systems neuroscience goal of understanding how neurons across an entire organism transform an environmental input to motor output. Recently, a confluence of advances has made such an investigion possible. The first is the advent of *C. elegans* as a powerful model organism.

1.1.1 *C. ELEGANS* AS A MODEL ORGANISM

Invertabrate model organisms have yielded many universal principles of neuroscience. Our understanding of the molecular basis of learning and memory, for example, came initially from studies of *Aplysia*, a giant sea slug. Similarly, our understanding of motor circuits was informed by early studies of leech and lamprey.

The nematode *Caenorhabditis elegans* has emerged as a popular and robust model organism and natural tool to study systems neuroscience. With only 302 neurons, its nervous system is compact but tractable. The 1 mm-long nematode exhibits a rich array of behaviors [33]. It senses its environment, navigates toward food [58] and temperatures [137] that it prefers, avoids chemicals that it dislikes [32], and responds to touch [22]. The worm even exhibits associative learning– it can associate odors with foods that make it sick, and avoid them accordingly, for example[175].

Starting in the 1970's, White et al. [160, 161] mapped the entire wiring diagram of the *C. elegans* nervous system. The individual neurons are morphologically identifiable and their connectivity is stereotyped from one worm to the next. The scientific community has therefore been able to systematically study individual neurons and use tools such as laser ablation and transgenics to identify which neurons are part of which neural circuits.

Critically, *C. elegans* is also a major platform for molecular genetics and genetic engineering. It was the first organism into which the fluorescent reporter gene GFP was cloned [23], and it was the first multicellular eukaryote to have its genome sequenced [1, 149]. The nematode is especially convenient for genetic engineering. The worms have a fast generation time (a worm grows from egg to egg-laying adult at room temperature in about four days). They are also naturally hermaphroditic and can self reproduce, making it trivial to maintain isogenic lines.

New genes can be added to the worm by injecting plasmids which are incorporated into an extrachromosomal array. The worms are then irradiated, and their DNA damage repair pathway incorporates the plasmids into their chromosomes to form stable transgenic lines. In the hands of a skilled geneticist, the whole process– including outcrossing –takes less than a month.

Importantly, *C. elegans* is also optically transparent, and thus particularly amenable to advances in optical physiology and microscopy, the primary tools used in this work.

These attributes have made *C. elegans* extremely well studied. As a result, there is a rich repository of knowledge and resources available to the *C. elegans* research community. Among the resources used commonly in the course of this work are: WormBase, an online database of genes, phenotypes and publications [65]; WormAtlas, a comprehensive online anatomical resource providing details of every neuron and cell [4]; WormBook, a curated collection of review articles and methods [151]; WormWeb, an online interactive network of neural connectivity [12]; and the *Caenorhabditis* Genetics Center at the University of Minnesota, which acts as a central repository and distributor for transgenic *C. elegans* lines.

C. elegans, with its compact nervous system, genetic tractability, optical access and well-mapped neural circuitry, is an ideal candidate for studying the neural dynamics underlying behavior.

1.1.2 Optical neurophysiology

Traditional electrophysiology is challenging to perform in the nematode *C. elegans* [57, 138]. The worm's small size and pressurized fluid-filled body make it difficult for electrodes to gain access to neurons. As a result, electrophysiology experiments are performed on worms that are partially dissected and immobilized. These preparations make it extremely difficult to correlate neural activity with behavior. Advances in optogenetics and fluorescent reporters, however, offer a viable alternative to electrophysiology. When used with transparent organisms, optical techniques are entirely non-invasive, making them the optimal choice for this work.

Optogenetics

Optogenetics is an emerging field that refers to optical tools based on genetically encoded proteins that manipulate neural function. Karl Deisseroth and Ed Boyden founded the field of optogenetics with their joint development of Channelrhodopsin as an optical method for neural stimulation in 2005 [14] and their simultaneous but independent development of Halorhodopsin as a method of neural silencing in 2007 [64, 174]. For a riveting historical account, see [13]. Optogenetic proteins like Channelrhodopsin and Halorhodopsin are light-activated transmembrane ion channels that open in response to light stimuli at a particular wavelength [48, 117, 171]. These optogenetic proteins were immediately employed in *C. elegans*, and an early experiment showed how illuminating worms expressing Channelrhodopsin in their mechanosensory neurons evoked a touch-like response [118]. The field has seen explosive growth in the past five years. For a review of optogenetics in *C. elegans*, including some of the work presented here, see [167] and [171]. This work uses optogenetics as a tool to probe neural activity in a freely moving worm.

Fluorescent Reporters of Neural Activity

Just as optogenetic proteins allow optical stimulation or inhibition of neural activity, genetically encoded fluorescent reporters provide optical readouts of neural activity. The first class of genetically encoded reporters were calcium indicators that altered their fluorescent properties in response to intracellular calcium levels. Calcium in a neuron is often used as a proxy for its membrane potential, and thus calcium indicators serve as an indirect measure of neural activity. The first genetically encoded calcium indicator was cameleon [115], a calmodulin protein modified by the addition of a Forster Resonance Energy Transfer (FRET) pair of fluorophores. As intracellular calcium levels increase, the calmodulin component of cameleon contracts and pulls the two attached fluorescing proteins closer together, changing their fluorescence properties. The past decade has seen steady improvement in genetically encoded calcium indicators that are brighter, more sensitive, and have greater dynamic range and faster response times [111, 116, 168]. In this work, I use one of the most recent indicators, GCaMP3 [153].

Both optogenetics and calcium indicators are non-invasive for transparent an-

imals. This, in principal, allows *C. elegans* to remain intact and unrestrained. Previously the Samuel Lab was the first to manually track a worm and observe its calcium transients as it moved freely [30]. In this thesis I develop an automated system to track the worm and either apply optogenetic stimuli or monitor calcium transients.

1.1.3 Real-Time Computer Vision

Conducting optical physiology on individual neurons in a moving worm requires significant ancillary hardware and software to identify the worm's outline, track it and keep it centered, identify targeted neurons within the worm and precisely illuminate and image those neurons in real-time –all while simultaneously recording the worm's macroscopic behavior. A number of technological advancements make this now possible. In addition to the development of optogenetics and fluorescent indicators, the ever decreasing cost of computer power, the availability of high power CW lasers in a variety of wavelengths, the advent of microelectromechanical systems (MEMS), and the development of powerful open source-computer vision libraries have conspired to make such a purely-optical system feasible for the first time.

It has only been in the past decade that the field of real-time computer vision algorithms has advanced to the point that libraries are now readily accessible to perform real-time video analysis using standard computer hardware. For example, this work utilizes hardware-optimized routines from the open-source OpenCV library [15, 16]. That library was only recently developed, in part, to address the DARPA Grand Challenge autonomous vehicle competition in 2004 and 2005 [143, 152]. Fortuitously, many closed-loop real-time image processing techniques for such applications can be applied to solve the optogenetic instrumentation problems of interest here. Consequently, this work builds upon prior computer vision developments from the DARPA Grand Challenge and elsewhere.

1.1.4 Applications to Behavioral Neuroscience

This work brings together advances in optogenetics and real-time computer vision to systematically perturb and monitor neural activity across an entire organism while it is behaving. At the time this thesis work began, it was the first successful

effort to bring these disparate strands together in the study of neural activity in *C. elegans*.

I have used the tools developed here to study neural activity driving *C. elegans* locomotion. The worm crawls on its side and propels itself forward by propagating bending waves from its head to tail. The worm has 95 muscles and 302 neurons at its disposal. Of those 302 neurons, 113 are classified as motor neurons, that directly enervate muscles. For successful locomotion, all of these muscles and neurons must work together in a coordinated and coherent fashion, interacting with sensory neurons and networks of neighboring neurons to process and respond to environmental conditions. For an excellent review of the *C. elegans* motor circuit see [156]. While it is fairly well understood which neurons in *C. elegans* are required for forward and backward locomotion, respectively, the dynamics of neural activity are largely unknown. In fact, many of the community's assumptions derive from leech and lamprey experiments [52, 82] and it is unclear to what extend those findings apply to *C. elegans*. Open questions include: Where are the oscillations that drive undulatory motion generated? Are the undulatory waves propagated through neurons or muscles? To what extent is feedback from the animal's environment required for locomotion?

Chapters 2 and 3 provide answers to some of these questions, and place strong experimental constrains on future models for *C. elegans* locomotion. Chapter 6 takes these questions to an even higher level, and asks how the worm's nervous system transitions between forward and backward locomotion.

1.1.5 Additional applications of computer vision for molecular biophysics

The computer vision techniques that I have developed here can also be applied to a wide variety of applications beyond behavioral neuroscience. In Chapters 4 and 5, I develop related software to automatically locate and identify fluorescing DNA origami nanorod barcodes . Unlike the chapters devoted to *C. elegans*, which focus primarily on the work in relation to *C. elegans* neuroscience, the chapters on DNA barcodes include rigorous descriptions of the software algorithms used and their mathematical underpinnings. All of the software described in this thesis is freely available under an open-source license. In general, the approach taken to

analyze images of DNA barcodes is representative of the approach to the computer vision problems in the remaining chapters. In particular, the number of steps in the algorithms in Chapter 5 offers a window into the complexity and challenge of quantitatively analyzing biological systems.

1.1.6 Onwards

Systems neuroscience is approaching a critical juncture. Arguably for the first time, the tools are becoming available to enable the complete characterization of how the neural activity of an entire organism drives behavior. Not only do the tools presented here allow for direct observation of neural activity, they also provide the means to perturb neural activity–both while simultaneously observing behavior. We now turn to the work itself.

2

Optogenetic manipulation of neural activity in freely moving *Caenorhabditis elegans*

2.1 Introduction

Reaserchers in systems neuroscience aim to understand how neural dynamics create behavior. Optogenetics has accelerated progress in this area by making it possible to stimulate or inhibit neurons that express light-activated proteins like channelrhodopsin-2 (ChR2) and halorhodopsin (also known as Halo/NpHR) by illuminating them [14, 27, 64, 117, 150, 173, 174]. The nematode *C. elegans* is particularly amenable to optogenetics due to its optical transparency, compact nervous system and ease of genetic manipulation [60, 98, 118, 146].

The ability to deliver light to one cell with spatial selectivity is essential for targeted optogenetic perturbation in *C. elegans* for the many cases where genetic methods do not provide adequate specificity. In the worm motor circuit, for ex-

ample, there are no known single neuron-specific promoters that would drive expression of light-activated proteins in only one or a few neurons of the ventral nerve cord (VNC). Optogenetics has been applied to the mechanosensory circuit in *C. elegans*, but only through simultaneous stimulation of all touch receptor neurons because promoters specific to each neuron are unavailable [117]. Researchers can use laser killing to study the contribution of single touch receptor neurons to overall behavior by removing neurons, but it is often preferable to work with intact circuits [22, 86, 162].

Recently, a digital micromirror device (DMD) has been used to deliver light with high spatial selectivity in immobilized *C. elegans* and immobilized *Danio rerio* zebrafish [164]. Each element of a DMD may be independently controlled to deliver light to a corresponding pixel of a microscope's field of view. In many cases, however, the normal operation of neural circuits can be studied only in freely behaving animals, requiring a more sophisticated instrument.

Here we describe an optogenetic illumination system that allows perturbations of neural activity with high spatial and temporal resolution in an unrestrained worm, enabling us to *Co*ntrol *L*ocomotion and *Be*havior in *R*eal *T*ime (CoLBeRT) in *C. elegans*. In the CoLBeRT system, a video camera follows a worm under dark-field illumination while a motorized stage keeps the worm centered in the camera's field of view. Machine-vision algorithms estimate the coordinates of targeted cells within the worm body and generate an illumination pattern that is projected onto the worm by a DMD with laser light. The cycle repeats itself for each subsequent frame. Because the worm is a moving target, the faster an image can be captured and translated into DMD directives, the more accurately an individual cell can be targeted. The CoLBeRT system carries out all of these functions in ~20 ms, providing a spatial resolution of ~30 μm in optogenetic control for freely swimming *C. elegans*. We analyzed the motor circuit and mechanosensory circuit of unrestrained worms, demonstrating the performance of the CoLBeRT system, a new tool that enhances the flexibility and power of optogenetic approaches in *C. elegans*.

2.2 Results

2.2.1 Experimental Setup

To stimulate neurons using ChR2 or inhibit neurons using Halo/NpHR, we used a 473-nm or 532-nm wavelength diode-pumped solid state (DPSS) laser, respectively (Fig. 2.1a). Either laser was incident onto a DMD with 1,024 × 768 elements. Laser light was reflected onto the specimen only when an individual micromirror was turned to the 'on' position. We illuminated the specimen under dark-field illumination by red light to avoid exciting ChR2 or Halo/NpHR. Filter cubes reflected the wavelengths for optogenetic illumination from the DMD onto the sample, while passing longer wavelengths for dark-field illumination to a camera. A motorized stage kept the specimen in the field of view.

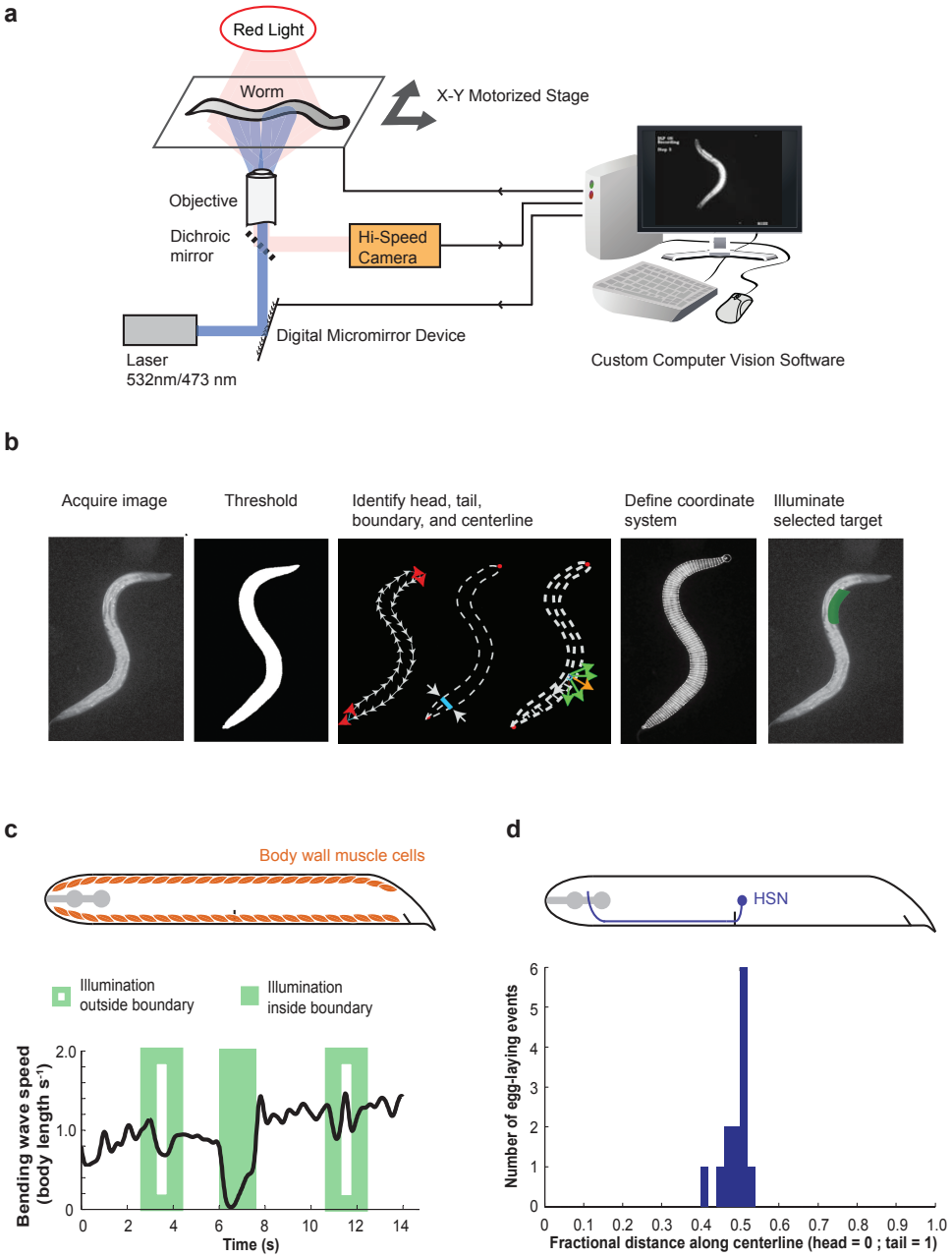
To accelerate real-time image analysis of worm posture, we developed the MindControl software package using the open-source OpenCV computer vision library[15]. With the graphical user interface (GUI), the user can dynamically target specific regions of freely moving worms. The MindControl software and documentation are available at http://github.com/samuellab/mindcontrol.

The MindControl software carries out a sequence of image analysis operations on each frame received from the camera (Fig. 2.1b). An image is captured by the computer, filtered and thresholded. Next, the boundary of the worm is calculated, and head and tail are identified as local maxima of boundary curvature (the head is blunt and the tail is sharp). The worm centerline is calculated and the body is divided into 100 evenly-spaced segments. These segments define a worm coordinate system invariant to worm posture or orientation, within which the user may define target positions. The software maps the position of targets onto the coordinates of the real image and, finally, sends the appropriate pattern to the DMD for illumination.

For our current system, the total latency between image acquisition and DMD illumination is 20 ms: image exposure, 2 ms; data transfer to computer, 3 ms; image analysis, 10 ms; and data transfer to DMD, 5 ms. Given the size and speed of a swimming worm at 10× magnification, our system working at ~50 frames per second (FPS) delivers optogenetic illumination with a spatial resolution of ~30 μm,

Figure 2.1 ***(following page)*****:** High-resolution optogenetic control of freely moving *C. elegans*. (**a**) An individual worm swims or crawls on a motorized stage under red dark-field illumination. A high-speed camera images the worm. Custom software instructs a DMD to reflect laser light onto targeted cells. (**b**) Images are acquired and processed at 50 FPS. Each 1,024 $\times$ 768 pixel image is thresholded and the worm boundary is found. Head and tail are located as maxima of boundary curvature (red arrows). Centerline is calculated from the midpoint of line segments connecting dorsal and ventral boundaries (blue bar) and is resampled to contain 100 equally-spaced points. The worm is partitioned into segments by finding vectors (green arrows) from centerline to boundary, and selecting one that is most perpendicular to the centerline (orange arrow). Targets defined in worm coordinates are transformed into image coordinates and sent to the DMD for illumination (green bar). (**c**) Schematic of body-wall muscles. Anterior, to left; dorsal, to top. Bending wave speed of swimming worm expressing Halo/NpHR in its body-wall muscles subjected to green light (10 mW mm^{-2}) outside or inside the worm boundary (n = 5 worms, representative trace). (**d**) Schematic of HSN. A swimming worm expressing ChR2 in HSN was subjected to blue light (5 mW mm^{-2}). Histogram, position at which egg-laying occurred when a narrow stripe of light was slowly scanned along the worm's centerline (n = 13 worms). Once an egg was laid, the worm was discarded.

Figure 2.1: (continued)



not far from the spatial resolution limit imposed by the pixel density of the DMD (~5 μm at 10× magnification).

2.2.2 Spatial resolution of the illumination system

First, we confirmed that illumination is restricted to the targeted area. We examined a transgenic worm expressing *Halo/NpHR::CFP* in all body-wall muscles. Whole-animal illumination of transgenic P*myo-3::Halo/NpHR* worms causes all muscles to relax [174]. We placed individual swimming worms in the CoLBeRT system and used green light (532 nm, 10 mW mm^{-2}) to alternately illuminate the entire region outside or inside the worm boundary (Fig. 2.1b and Video 2.6.1). Illuminating the entire region outside the worm boundary had no effect– the bending waves propagated from head to tail at normal speed. Illuminating the entire region inside the worm boundary, however, arrested locomotion— the body relaxed and the speed of bending waves dropped to zero.

To quantify the spatial resolution of the CoLBeRT system, we measured its targeting accuracy in evoking egg-laying events by stimulating the HSN motor neurons. We used transgenic worms expressing ChR2 under the *egl-6* promoter, which drives expression in the bilaterally symmetric HSN neurons (HSNL and HSNR) as well as glia-like cells in the worm's head [133]. Optogenetic stimulation of the HSN neurons, which innervate the vulval musculature, evokes egg-laying behavior (L. Emtage and N. Ringstad, personal communication).

The two HSN neurons lie on top of one another when the worm is viewed laterally, so our system targets both neurons. We projected a thin stripe of blue light (473 nm, 5 mW mm^{-2}) on the body of swimming P*egl-6d::ChR2* transgenic worms. The long axis of the stripe was orthogonal to the worm centerline and spanned its diameter. The stripe width corresponded to 2% of the anterior-posterior length of the worm centerline (that is, ~20 μm of the ~1-mm-long young adult worm). We used narrow stripes so that our illumination would less likely stimulate HSN when illuminating its process. We slowly moved the illumination stripe along the centerline of swimming worms while recording egg-laying events. Of 14 worms studied, we observed 13 egg-laying events, eight in which the stripe started at the head and five in which the stripe started at the tail. Egg-laying frequency sharply peaked when the center of the stripe coincided with the centerline coordinate of

the HSN cell bodies, or 49.6% of the total distance from the anterior to the posterior of the body with 3.2% s.d. (Fig. 2.1d and Video 2.6.2). The width of this distribution suggests that the CoLBeRT system provides at least ~30 μm of spatial resolution.

2.2.3 Optogenetic manipulation of muscle cells

In *C. elegans*, forward movement is driven by motor neurons in the VNC, which coordinate the activity of 95 body wall muscle cells along the dorsal and ventral sides of the VNC [156]. The circuit for worm locomotion is poorly understood in comparison to that of other undulatory animals such as the leech and lamprey [18, 82, 113]. Because this circuit probably operates normally only during normal movement, technology such as the CoLBeRT system is necessary to dissect cellular activity in unrestrained animals.

We used the CoLBeRT system to suppress muscle activity in a region of the body in P*myo*-3 ::*Halo/NpHR*::*CFP* transgenic worms (Fig. 2.2 and Video 2.6.3). This perturbation of undulatory dynamics can be shown graphically using a red-blue color map to represent the curvature of the body centerline in nondimensional units (that is, the curvature calculated at each point along the centerline, κ, multiplied by worm length, L) as a function of time and fractional distance along the centerline, s, from head ($s = 0$) to tail ($s = 1$) (Fig. 2.2a). Notably, hyperpolarizing muscle cells in one segment had no effect on undulatory dynamics anterior to the segment, but lowered the amplitude of the bending wave posterior to the illuminated segment (Fig. 2.2b). Representative data from one of five worms that we studied are shown in Figure 2.2. Thus, the bending of posterior body segments seems coupled to the bending of anterior body segments. One possibility is that muscle activity in posterior segments is directly promoted by muscle activity in anterior segments, perhaps by gap junction coupling between muscle cells [103]. Another possibility is that the motor circuit contains a proprioceptive mechanism that makes the activity of posterior segments directly sensitive to the bending of anterior segments.

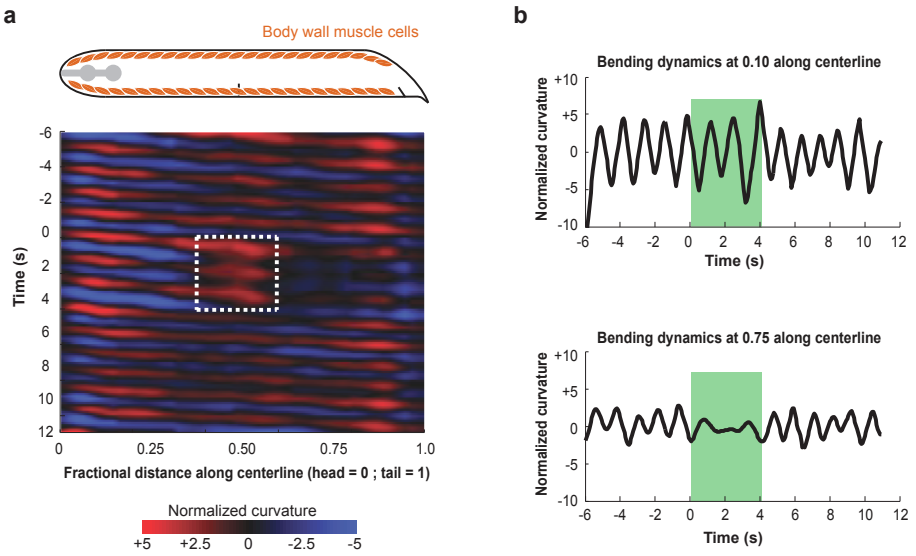


Figure 2.2: Optogenetic inactivation of muscle cells. (**a**) Kymograph of time-varying body curvature along the centerline of a P*myo-3::Halo/NpHR::CFP* transgenic worm. Between 0 s and 4 s, the worm was illuminated with green light (10 mW mm^{-2}) in a region spanning the worm diameter and between 0.38 and 0.6 of the fractional distance along the centerline. (**b**) For the kymograph in **a**, time-varying curvature at two points along the worm centerline, both anterior (top) and posterior (bottom) to the illuminated region.

2.2.4 Optogenetic manipulation of cholinergic motor neurons

The cell bodies of motor neurons in *C. elegans* are distributed along the VNC [156, 162]. Ventral muscles are innervated by the cholinergic VA, VB and VC motor neurons and GABAergic VD motor neurons. Dorsal muscles are innervated by the cholinergic DA, DB and AS motor neurons and GABAergic DD motor neurons[25, 160]. A current model is that VA and DA drive muscle contraction during backward locomotion, VB and DB drive muscle contraction during forward locomotion and VD and DD motor neurons drive muscle relaxation during both forward and backward locomotion [67, 156, 162]. A repeating motif of synaptic connectivity between the motor neurons runs along the worm body and allows for contralateral inhibition [156]. During forward locomotion, for example, the DB (or VB) motor neurons can simultaneously excite a dorsal (or ventral) muscle cell while exciting the GABAergic VD (or DD) motor neurons that inhibit the opposing ventral (or dorsal) muscle cell [25, 160]. How this network drives the rhythmic undulatory wave, however, is poorly understood.

We analyzed the contributions of cholinergic neurons to forward locomotion using transgenic worms expressing Halo/NpHR in all cholinergic neurons under the control of the *unc-17* promoter [134]. In P*unc-17::Halo/NpHR::CFP* transgenic worms, illumination of a short segment of the VNC suppressed propagation of the undulatory wave to the entire region posterior to the illuminated segment without affecting the undulatory wave anterior to the illuminated segment (Fig. 2.3a,b and Video 2.6.4). Representative data from one of five worms that we studied are shown in Figure 2.3a,b. This suggests that the activity of posterior VB and DB neurons is coupled to the activity of anterior VB and DB neurons, consistent with a wave of neuronal excitation that propagates from head to tail during forward movement.

Using the CoLBeRT system, we can also specifically illuminate either the dorsal nerve cord or the VNC (Video 2.6.5). The VNC contains the cell bodies of the cholinergic motor neurons, whereas the dorsal nerve cord contains only nerve processes. Illuminating the entire VNC was particularly effective in hyperpolarizing the cholinergic motor neurons of P*unc-17::Halo/NpHR::CFP* worms, inducing paralysis. Illuminating the entire dorsal nerve cord, however, produced only a

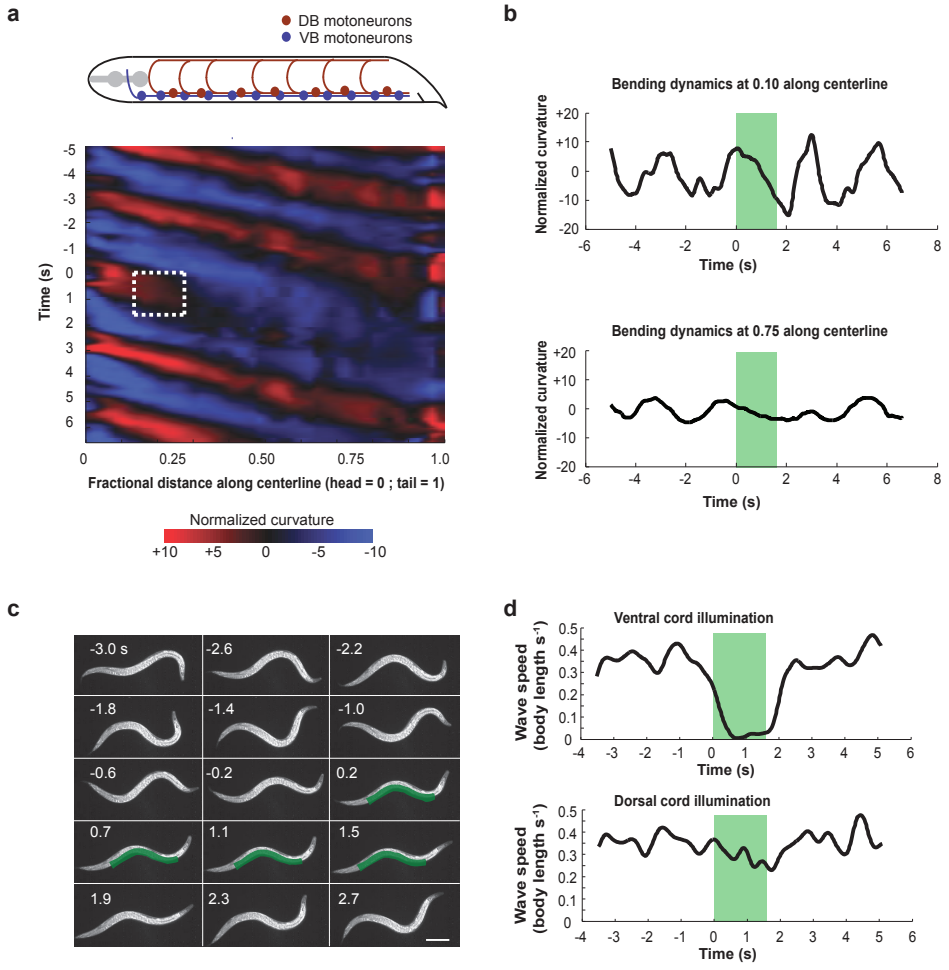


Figure 2.3: Inhibition of motor neurons. (**a**) Schematic of cholinergic DB and VB motor neurons. Anterior, to left; dorsal, to top. Kymograph of time-varying body curvature along the centerline of a P*unc-17::Halo/NpHR::CFP* transgenic worm illuminated by a stripe of green light (10 mW mm^{-2}) along its VNC between $t = 0$ s and 1.6 s. In the dorsal-ventral direction, the stripe width was equal to 50% of the worm diameter and centered on the ventral boundary. In the anterior-posterior direction, the stripe length was between 0.14 and 0.28 of the fractional distance along the body. (**b**) For the kymograph in a, time-varying curvature at two points along the worm centerline, both anterior (top) and posterior (bottom) to the illuminated region. (**c**) Video sequence of worm illuminated by a long stripe of green light (10 mW mm^{-2}) spanning the VNC between $t = 0$ s and 1.8 s. Scale bar, 100 µm. (**d**) Bending wave speed of a swimming worm illuminated by a long stripe of green light (10 mW mm^{-2}) lasting 1.8 s and spanning the VNC (top) and dorsal nerve cord (bottom).

small (~15%) drop in the speed of wave propagation (Fig. 2.3c,d). The asymmetric effect of illuminating the ventral and dorsal nerve cords is probably due to the higher density of optogenetic protein in the cell bodies.

Surprisingly, the paralysis evoked by illuminating the VNC can occur without allowing relaxation of the worm body. In this instance, as long as the entire cholinergic network within the VNC was deactivated, the worm retained the posture it had immediately before illumination (Fig. 2.3c). When the muscle cells of a swimming worm were hyperpolarized, on the other hand, the body straightened (Video 2.6.1). This observation suggests that muscle cells can remain in contracted or relaxed states without requiring continuous cholinergic input.

2.2.5 Optogenetic manipulation of single touch receptor types

Next, we applied the CoLBeRT system to the touch receptor system in *C. elegans*. Six cells are specialized for sensing gentle touch in *C. elegans*: the left and right anterior lateral microtubule cells (ALML and ALMR, respectively); the left and right posterior lateral microtubule cells (PLML and PLMR, respectively); the anterior ventral microtubule cell (AVM); and the posterior ventral microtubule cell (PVM) [22]. Gently touching the worm near its anterior stimulates reversal movement dependent on ALML, ALMR and AVM. Gently touching the worm near its posterior stimulates forward movement dependent on PLML and PLMR. The role of PVM remains unclear.

Channelrhodopsin can be expressed in all six touch receptor cells using the *mec-4* promoter. Illuminating the whole body of transgenic worms with blue light evokes reversal responses, presumably by simultaneously activating ALM, AVM and PLM1. With the spatial resolution afforded by the CoLBeRT system, we could individually activate the ALM, AVM and PLM cell types. The left and right lateral cells (ALML and ALMR; PLML and PLMR) lie on top of one another when the worm is viewed laterally. Illuminating the anterior end containing both the AVM and ALM neurons triggered reverse movement (Fig. 2.4a and Video 2.6.6). Illuminating the posterior end containing the PLM neurons triggered forward movement (Fig. 2.4b and Video 2.6.7). Representative data from one of five worms that

we studied are shown in Figure 2.4a,b.

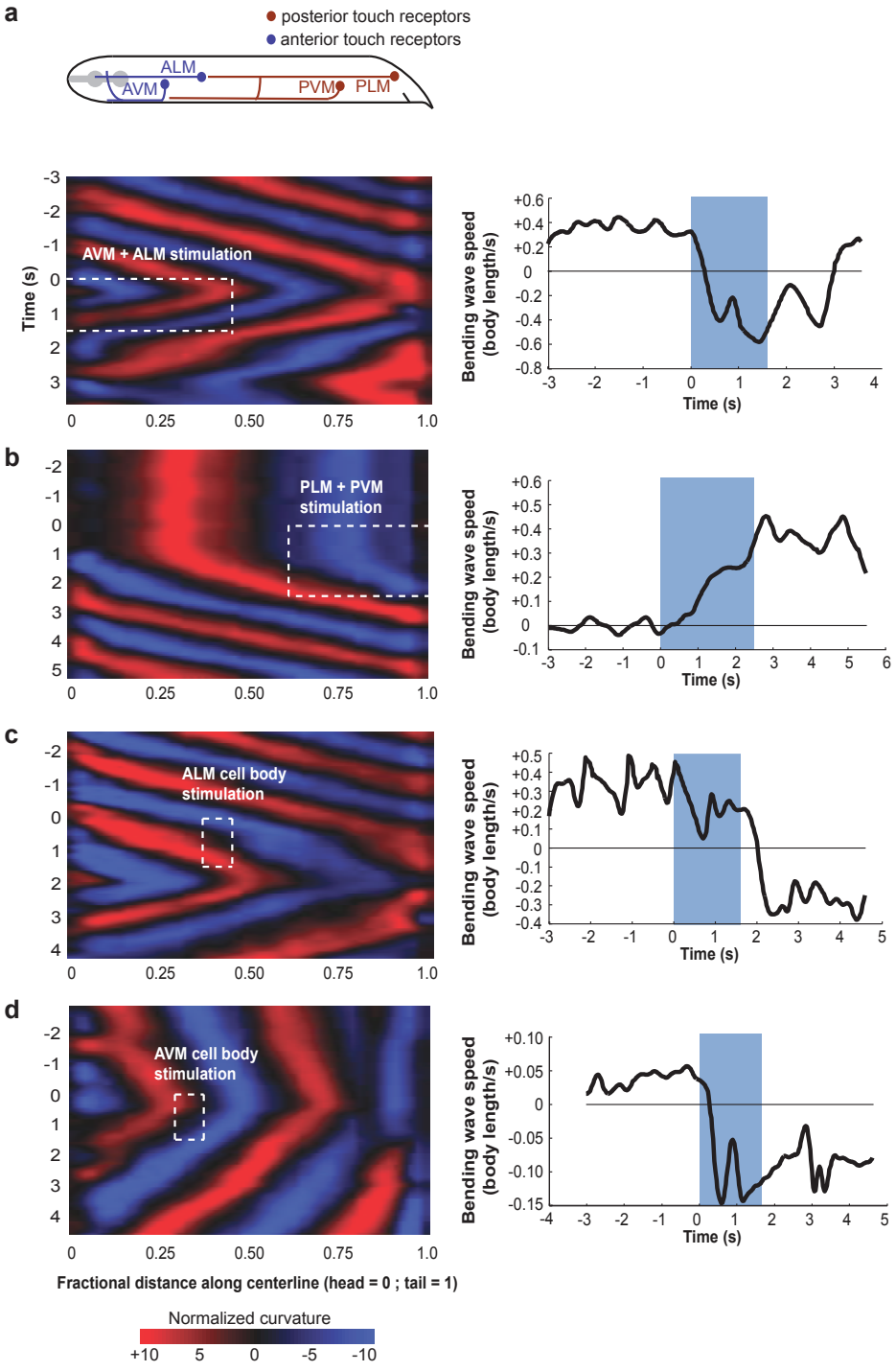
Using the CoLBeRT system, we also induced reversals by targeting just AVM or ALM with an illumination box (20 μm in the dorsal-ventral dimension; 30 μm in the anterior-posterior direction for a young adult worm) that was centered on each cell body (Fig. 2.4c,d and Videos 2.6.8 and 2.6.9). Representative data from one of 14 worms that we studied are shown in Figure 2.4c,d. Using these illumination boxes, we could avoid illuminating the axon of the nontargeted neuron. These observations are consistent with earlier work showing that single touch receptor types are sufficient to drive behavioral responses [21].

To confirm that the CoLBeRT system can specifically target either AVM or ALM, we used transgenic worms expressing the photoconvertible fluorescent protein Kaede in the mechanosensory neurons [6]. Upon illumination by UV or violet light, Kaede converts from a green to red fluorescent state. We used the CoLBeRT system with 405-nm light to specifically illuminate either the AVM or ALM cell bodies for 60 s in freely moving P*mec-4::Kaede* worms. We found that worms in which AVM or ALM had been targeted showed only detectable red fluorescence in AVM or ALM, respectively, whereas all mechanosensory neurons showed green fluorescence (Fig. 2.5a,b). When targeting ALM, a transient segmentation error owing to an omega turn by the worm caused the system to illuminate PLM and PVM for 1 s, producing slight photoconversion in those neurons (Fig. 2.5b). By quantifying the ratio between the red and green fluorescence signals, we estimated that the nontargeted neurons were illuminated for less than ~1 s (Methods).

It has been shown that the mechanosensory circuit habituates to repetitive optogenetic stimulation [118]. We used the CoLBeRT system to quantify the rate of AVM and ALM habituation over 40 minuntes by repeatedly stimulating either AVM or ALM every 60 s. We observed comparable rates of habituation for both ALM and AVM (Fig. 2.5c,d). Others have studied loci for habituation in the mechanosensory circuit by laser-killing touch receptor cells and/or downstream neurons and quantifying rates of habituation to gentle touch [164]. If habituation partly occurs at interneurons that are downstream of both ALM and AVM, then we might expect cross-habituation of the AVM response to repeated ALM stimulation, and vice versa. Cross-habituation may also be mediated by an electrical gap junction between AVM and ALM [160]. To test whether cross-habituation

Figure 2.4 ***(following page)*****:** Optogenetic analysis of mechanosensory neurons. (**a**)Top, schematic of anterior and posterior touch receptor cells. Anterior, to left; dorsal, to top. Kymographs (left) of time-varying curvature of centerline of worms expressing ChR2 in mechanosensory neurons (P*mec-4::ChR2::GFP*) subjected to rectangles of blue light (5 mW mm^{-2}) targeting different groups of touch receptor neurons. Plots of bending wave speed (right) indicate stimulus-evoked changes in direction or speed. AVM and ALM neurons are subjected to 1.5 s of stimulation. Given a coordinate system where x specifies dorsal-ventral location (–1, dorsal boundary; 0, centerline; 1, ventral boundary) and y defines fractional distance along the worm's centerline (0, head; 1, tail), the rectangle of illumination has corners (x,y) = ((-1.1,0),(1.1,0.46)). (**b**) PVM and PLM neurons are subjected to 2.5 s of stimulation with a rectangular illumination ($n = 5$ worms, representative trace) with corners at (x,y) = ((-1.1,0.62),(1.1,0.99)). (**c**) ALM cell body is specifically stimulated by illuminating a small rectangle with corners at (x,y) = ((-0.3,0.38), (-0.9,0.46)). (**d**) AVM cell body is specifically stimulated by illuminating a small rectangle with corners at (x,y) = ((0.3,0.3),(0.9,0.38)).

Figure 2.4: (continued)



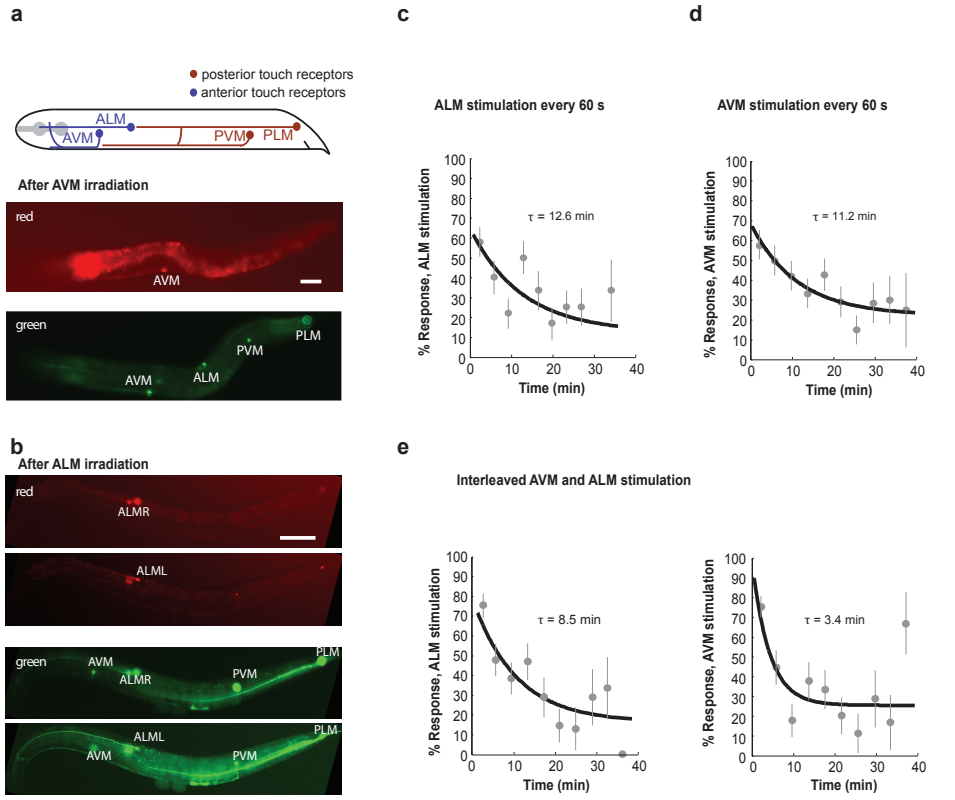


Figure 2.5: Habituation of individual touch receptor neuronal types. (**a,b**) Schematic showing anterior and posterior touch receptor neurons (top). Anterior, to left; dorsal, to top. A freely swimming worm expressing Kaede in touch receptor neurons was continuously tracked and illuminated with a small rectangle of 405 nm light (2 mW mm^{-2}) centered on either AVM or ALM (as in Fig. 2.4c,d) for 60 s. Red and green fluorescence images are shown. Scale bars, 100 µm. (**c–e**) Individual ALM and AVM neurons were repeatedly stimulated with blue light (5 mW mm^{-2}) for 1.5 s every 60 s for ~40 min, either alone (**c,d**) or interleaved within each experiment (**e**; ALM, 30 s; AVM, 30 s; ALM, 30 s; and so on). Fractional response to stimulus of each neuronal type was fit to an exponential, $a + b\exp(t/\tau)$, using maximum likelihood estimator. Time constant for habituation, τ, was extracted from each fit. Error bars, s.e.m. Fractional response of ALM when stimulated alone (**c**; $n = 7$ worms). Fractional response of AVM when stimulated alone (**d**; $n = 8$ worms). Fractional response of ALM (left) and AVM (right) during interleaved stimulation of both (**e** ; $n = 7$ worms).

occurs, we subjected a worm to interleaved AVM and ALM stimulation every 30 s, such that each neuron type was stimulated every 60 s. We found that the rates of habituation to both AVM and ALM stimulation were indeed more rapid with interleaved stimulation than with individual stimulation. This effect was particularly pronounced in the case of AVM stimulation (Fig. 2.5e).

2.3 Discussion

At present, the spatial resolution of CoLBeRT is ~30 μm when tracking a swimming worm. The system has better resolution when tracking the slower movements of a crawling worm, but is ultimately limited to ~5 μm resolution due to the pixel resolution of the DMD. Higher spatial resolution could be reached by tracking a specific region of the worm (for example, the nerve ring) at higher magnification. This modification to CoLBeRT would require a different approach to image analysis and targeting– for example, analysis of cell body fluorescence instead of analysis of the posture of the whole worm.

CoLBeRT may be adapted to the optogenetic analysis of other genetically tractable, transparent animals such as the larvae of *Drosophila melanogaster* or *D. rerio*. A simplified version of CoLBeRT may also be used to facilitate optogenetic illumination in other settings like studies of mammalian brain slices or exposed brain surfaces. Variants of CoLBeRT, using its capacity for rapid closed-loop feedback, may be used to trigger optogenetic stimulation based on simultaneous recordings of neural activity in addition to animal posture.

CoLBeRT is a flexible and easy-to-use platform for designing and projecting arbitrary spatiotemporal patterns of illumination with closed-loop sensitivity to the real-time behavior of the worm.

2.4 Methods

2.4.1 Strains

We cultivated transgenic worms in the dark at 20 °C on nematode growth medium (NGM) plates with OP50 bacteria with all-trans retinal. We made OP50-retinal plates by seeding 6-cm NGM plates with 250 μl of a suspension of OP50 bacteria

in LB, to which we added 1 μl of 100 mM retinal in ethanol immediately before seeding. Plates were stored in the dark and all worms were handled in the dark or under red light.

Strain FQ10 (P*egl-6::ChR2::YFP*) was a gift of Nials Ringstad. Strain QH3341 (vdEx128(P*mec-4::Kaede*)) was a gift of Brett Neumann and Massimo Hilliard. Strains ZX444 (*lin-15*(n765ts); zxEx29 (P*myo-3::NpHR::ECFP*; *lin-15*+)) and ZX422 (*lin-15*(n765ts); zxEx33 (P*unc-17::NpHR::ECFP*; *lin-15*+)) were gifts of Alexander Gottschalk. Strain P*myo-3::Halo::CFP* used in our experiments was generated by integrating the transgene in ZX444 by cobalt-60 irradiation and outcrossing the resulting strain three times to the wild-type N2 strain. Strain P*unc-17::Halo::CFP* used in our experiments was generated by Mei Zhen by irradiating ZX422 using UV radiation and outcrossing twice to the wild-type N2 strain. The P*mec-4::ChR2* strain (QW309) was generated by injection of P*mec-4::ChR2::YFP* plasmid at 100 ng μl^{-1} into *lin-15*(n765ts) worms along with the *lin-15* rescuing plasmid (pL15 EK) at 50 ng μl^{-1}. The extrachromosomal array was integrated using gamma irradiation and outcrossed four times to wild-type N2.

2.4.2 Microscopy

The setup was built around a Nikon Eclipse TE2000-U inverted microscope. We carried out dark-field imaging using annular illumination of the specimen through a Ph3 phase ring. A filter transmitting red light (Hoya) was mounted to the microscope illumination optical pathway to minimize inadvertent activation of ChR2 or Halo/NpHR owing to dark-field illumination.

We imaged worms using a 10×, numerical aperture (NA) 0.45 Plan Apo objective. We used a custom optical system composed of two camera lenses (Nikon) to reduce the size of the image on the camera by a factor of 3.5. This allowed us to capture almost all of the 2.5-mm-diameter field of view on the camera sensor. We used a PhotonFocus MV2-D1280-640CL camera and BitFlow Karbon PCI Express ×8 10-tap Full Camera Link frame grabber to capture images.

The microscope stage was controlled by a Ludl BioPrecision2 XY motorized stage and MAC 6000 stage-controller. During data acquisition, computer software kept the worm centered in the field of view via an automated feedback loop.

2.4.3 Optics and illumination

To stimulate ChR2, we used a DPSS laser (LP473-100, 473-nm wavelength, 100-mW maximum power, LaserShowParts). Similarly, to stimulate Halo/NpHR we used a DPSS laser (LP532-200, 532-nm wavelength, 200-mW maximum power, LaserShowParts). To photoconvert Kaede, we used a DPSS laser (EL-100B, 405-nm wavelength, 100-mW maximum power, Laserwold). The beams from the 473-nm and 532-nm lasers were aligned to a common path by a dichroic beamsplitter. The beam from the 405-nm laser was aligned to the common path with a retractable mirror. For each experiment, however, only one of the three lasers was used. The laser beam was expanded using a telescope composed of two plano-convex lenses and incident onto a 1,024 × 768 pixel digital micromirror device (Texas Instruments DLP, Discovery 4000 BD VIS 0.55-inch XGA, Digital Light Innovations) attached to a mirror mount. Using a series of mirrors, the laser was aligned so that the reflected beam for the 'on' state of the DMD was centered on the optical axis of the illumination pathway.

The plane of the DMD was imaged onto the sample via the epifluorescence illumination pathway of the microscope using an optical system composed of two achromatic doublet lenses. We used a dichroic filter, FF580-FDi01-25x36 (Semrock), to reflect 405-nm, 473-nm or 532-nm laser light onto the sample while passing wavelengths used for dark-field illumination (λ > 600 nm). We used an emission filter, BLP01-594R-25 (Semrock), to prevent stray laser reflections from reaching the camera. The dichroic and emission filters were mounted in a custom filter cube in the microscope filter turret.

2.4.4 Software

The microscope and all its components were controlled with custom MindControl software running Windows XP on an Acer Veriton M670G computer with an Intel Core 2 Quad processor running at 2.83 GHz with 3 GB of RAM. MindControl enables the user to define arbitrary illumination patterns for optogenetic stimulation, and to deliver illumination patterns either manually or automatically. For rapid operation, MindControl was written in the C programming language using the open-source OpenCV computer vision library, along with Intel's Integrated Performance Primitives for maximal speed. To further increase speed, we used

multiple threads to separately handle image processing and the user interface. Every 20 ms, MindControl acquires an image from the camera, computes the location of the worm, generates an illumination pattern and sends that pattern to the DMD.

For each video frame, the boundary and centerline of the worm and the status of the stimulus is recorded in a human- and computer-readable YAML file. Every frame is also recorded in two video streams, one containing annotations about optogenetic stimulation, and the other containing only images of the freely moving worm. A GUI allows the user to adjust the parameters of optogenetic stimulation in real time during each experiment. After each experiment, we used custom scripts written in Matlab to carry out quantitative analysis of the resulting video. All software and documentation is freely available for modification and redistribution under the GNU general public license. The software for optogenetic stimulis and analysis are available at `http://github.com/samuellab/mindcontrol` and `http://github.com/samuellab/mindcontrol-analysis`, respectively.

2.4.5 Behavioral experiments

For motor circuit experiments, we washed each young adult worm in NGM solution before transferring it to a chamber composed of ~100 μl of a 30% dextran (wt/vol) in NGM solution sandwiched between two microscope slides separated by 0.127 mm. In this chamber, the worm was approximately confined to two dimensions but otherwise able to move freely. We then placed the chamber on the microscope for data collection.

To analyze egg-laying, we selected gravid adult worms, washed them in NGM and transferred them to chambers as described above. Each worm was subject to sequential pulses of 4 s of blue light illumination. Each pulse illuminated a stripe orthogonal to the worm centerline, spanning the worm diameter with width corresponding to 2% of total body length. The stripe progressed along the worm centerline from head to tail or from tail to head until the first egg was laid. After an egg was laid, the trial ended and the worm was killed. Out of 14 worms studied, one did not lay any eggs.

For mechanosensory circuit experiments, we prescreened young adult P*mec-4::ChR2* worms on a fluorescence stereo microscope (Nikon SMZ 1500) by illumi-

nating the anterior of the worm with blue light from a 50-W mercury lamp through a GFP excitation filter. Only worms that responded with a reversal were chosen for further experiments. We carried out this prescreening procedure because the P*mec-4::ChR2* strain (QW309) showed noticeable worm-to-worm variability. Only ~70% responded robustly and consistently. The reasons for this variability are unclear. Worms that passed this prescreening were then transferred to an unseeded NGM agar plate and allowed to crawl for ~30 s to free themselves of bacteria.

We then transferred worms onto a plate containing a 1 to 2-mm-thick layer of NGM agar and covered with mineral oil to improve optical imaging quality. Specific regions of each worm were targeted with blue light and illuminated for 1.5 s. We scored anterior touch responses by quantifying the bending wave speed 2 s before stimulus onset and 3 s after stimulus onset. We classified a successful response to stimuli as a reduction in wave speed by > 0.03 body lengths per second. To calculate habituation rates, as in Figure 2.5c–e, multiple worms were repeatedly stimulated over time. Fractional response, as plotted, is the total number of observed responses divided by the total number of stimuli in a ~4 min window for all worms in a given experiment.

2.4.6 Quantifying locomotory behavior

The locomotory behavior of individual worms was analyzed by quantifying time-varying worm posture in each video sequence. A least-squares cubic smoothing spline fit to the body centerline was calculated. Curvature was calculated at each point along the centerline as the derivative of the unit vector tangent to the centerline with respect to the distance along the centerline. To graphically display locomotory gait, we use kymographs of curvature as a function of distance along the centerline and time. We calculated the speed of the bending wave along the centerline within the reference frame of the worm body by measuring the displacement of curvature profiles along the centerline (Δx) at successive points in time (Δt) according to $\nu = \Delta x / \Delta t$.

2.5 Acknowledgements

This work was supported by the Dana Foundation, the US National Science Foundation and a US National Insitutes of Health Pioneer Award to Aravinthan

D.T. Samuel. Andrew M. Leifer is supported by National Science Foundation Graduate Research fellowship Grant No. DGE-0644491. We thank Mei Zhen (Samuel Lunenfeld Institute), Niels Ringstad (Skirball Institute of Biomolecular Medicine, New York University School of Medicine), Alexander Gottschalk (Frankfurt Molecular Life Sciences Institute) and Brent Neumann and Massimo Hilliard (Queensland Brain Institute, University of Queens) for gifts of transgenic strains; Jeffrey Stirman for sharing unpublished results about a similar system that he developed; Brian Chow and Theodore Lindsay for useful discussions; Anji Tang and Benjamin Schwartz for assistance with data analysis; and Christopher Clark for making the *mec-4* transgenic worm.

2.6 Accompanying Videos

2.6.1 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S2.mov` (5 MB)

A P*myo-3::Halo::CFP* worm expressing Halorhodopsin in muscle is induced to relax only when the CoLBeRT system illuminates within the worm's body. The movie shows the same individual as shown in Figure 2.1c. During frames 6707 to 6771, the entire region outside the worm's boundary is illuminated with green light (10 mW mm^2) and the worm continues locomotion. During frames 6,847 to 6,917, only the region inside the boundary of the worm is illuminated and the worm relaxes. During frames 7,052 to 7,117 only the region outside the worm's boundary is illuminated and the worm continues moving normally. The frame number is indicated at the bottom right. Light green shading indicates the area the system is targeting. Bright green shading and the appearance of the words "DLP ON" indicate that the system is illuminating the targeted area.

2.6.2 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S3.mov` (8 MB) An P*egl-6::ChR2::GFP* worm is induced to lay eggs when a stripe of blue light reaches HSN. The video shows the same individual as in Figure 2.1d. A narrow stripe of light (5 mW mm^2), 0.02 of the fractional length along the

worm centerline and twice the width of the worm, progresses from the worm's head toward its tail. The stripe takes steps of 0.02 fractional worm lengths and illuminates for 4 s at each step. At frame 8,828, the illumination band reaches HSN and the worm lays eggs. The frame number is indicated at the bottom right.

2.6.3 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S4.mov` (960 kB) The bending waves of a P*myo-3::Halo::CFP* transgenic worm are dampened and the anterior relaxes when a portion of the worm is illuminated with green light. The video shows the same individual as in Figure 2.2. The illumination is turned on 4 s into the video. The worm recovers after the illumination is turned off. Light green shading indicates the area the system is targeting. Bright green shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.6.4 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S5.mov` (3 MB) The bending waves of an P*unc-17::Halo::CFP* are abolished when a small ventral region near the worm's head is illuminated. The video shows the same individual as shown in Figure 2.3a,b. During frames 9,075 to 9,141, the worm is illuminated with green light (10 mW mm^2) and no bending waves are propagated from the head to the tail. On the contrary, the worm is paralyzed posterior to the region of illumination and its curvature is frozen. Only after the stimulation ends are bending waves again able to propagate from the anterior to posterior of the worm. The frame number is indicated at the bottom right. Light green shading indicates the area the system is targeting. Bright green shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.6.5 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S6.qt` (5 MB) An P*unc-17::Halo::CFP* transgenic worm is paralyzed only when the

ventral nerve cord is illuminated, but not when the dorsal nerve cord is illuminated. The video shows the same individual as in Figure 2.3c,d. The ventral nerve cord is illuminated with green light at 10 mW mm² (frames 37,909 to 37,971) and then the the dorsal nerve cord is illuminated (frames 38,233 to 38,295). Note that during paralysis the worm does not relax to a neutral position. Light green shading indicates the area the system is targeting. Bright green shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.6.6 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S7.mov` (988 kb) The anterior of a P*mec-4::ChR2::GFP* worm is illuminated for 1.5 s, inducing a reversal. The video shows the same individual as in Figure 4a. During frames 7,645 to 7,709, the anterior 46% of the worm is illuminated with blue light at 5 mW mm², which includes the neurons AVM and ALM and their associated processes. The frame number is indicated in the bottom right hand corner. Light blue shading indicates the area the system is targeting. Bright blue shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.6.7 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S8.mov` (8 MB)

The posterior of a P*mec-4::ChR2::GFP* worm is illuminated with blue light, inducing forward movement. The video shows the same individual as in Figure 4b. During frames 13,655 to 13,733, the posterior 38% of the worm–which includes the neurons PVM and PLM and their associated processes–is illuminated with blue light (5 mW mm²) for 1.5 s. The worm, initially in a resting state, moves forward. The frame number is indicated in the bottom right hand corner. Light blue shading indicates the area the system is targeting. Bright blue shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.6.8 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S10.mov` (2 MB)

The cell bodies of ALM in a P*mec-4::ChR2::GFP* worm are illuminated with blue light, inducing a reversal. The video shows the same individual as in Figure 2.4c. During frames 2,013 to 2,079, ALM is illuminated with blue light (5 mW mm^2) for 1.5 s. The worm subsequently reverses. The frame number is indicated in the bottom right hand corner. Light blue shading indicates the area the system is targeting. Bright blue shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.6.9 Video

`http://www.nature.com/nmeth/journal/v8/n2/extref/nmeth.1554-S10.mov` (896 kB)

The cell body of the single neuron AVM in a P*mec-4::ChR2::GFP* worm is illuminated with blue light, initiating a reversal. The video shows the same individual as in Figure 2.4d. During frames 1,925 to 1,994, AVM is illuminated with blue light (5 mW mm^2) for 1.5 s and the worm subsequently undergoes a reversal. The frame number is indicated in the bottom right hand corner. Light blue shading indicates the area the system is targeting. Bright blue shading and the appearance of the words "DLP ON" indicate that the system is illuminating the target.

2.7 Design Considerations

The CoLBeRT system demands a delicate balancing of competing design criteria and pushes the limits of speed and computational power on multiple fronts. Moreover, the CoLBeRT is the first of its kind, so at the time of its design were no existing templates upon which to follow. In this section I discuss tradeoffs and decisions that informed the CoLBeRT system's development.

2.7.1 Guiding Principles

When developing the system, the following principles served as a guide:

1. *Low latency is crucial.* The accuracy of the CoLBeRT system depends on the time that elapses between imaging the worm and using that image to aim the laser light. Every component and system was designed to minimize latency. Note that latency is not the same as the inverse of the system's frame rate.

2. *Eschew obfuscation.* There are often hidden costs to complexity. Wherever possible I pursued the simplest available solution.

3. *Iterate and modularize.* Complicated tasks were broken into small ones. For example, before building the CoLBeRT system to work with microscopic worms, I first built a prototype version that projected light onto targeted regions of 8½" by 11" photographs of worms.[1]

4. *Promote rapid prototyping.* Scientific objectives and experimental conditions are apt to change. I chose flexible technologies that enabled rapid prototyping over technologies that may have been more optimal but also more brittle.

5. *Maintain transparency in software, data and development.* The software, instruments and experiments are complicated and produce large quantities of data. Where possible, I have sought out systems that make it easier to understand what a given piece of code does or what a given snippet of data means and when and under what conditions that data was created.

In the following I highlight how these principles have guided CoLBeRT's design.

2.7.2 Choosing a DMD

A DMD Can be obtained by either modifying an off-the-shelf projector, or purchasing a commercial DMD developer's kit directly from a Texas Instruments distributor. Initially, using an off-the-shelf projector appeared attractive. Off-the-shelf projectors are inexpensive and can interface to a computer over a simple VGA cable, whereas the developer's kit requires writing custom software to interface with the DMD controller over USB. I decided on using a developer's kit (see Section 2.4.3) for a number of reasons.

[1] For a video of the early tabletop prototype, see `http://vimeo.com/18840631`.

First, an off-the-shelf projector has additional unnecessary complexity. Any off-the-shelf projector is designed to project color images and thus employs considerable extraneous machinery. For example, in an off-the-shelf projector the DMD adjusts its mirrors multiple times within a frame so as to shine three temporally distinct images, one each under a red, green and blue filter. This is often synchronized to a spinning filter wheel that lies in the path of the projected light. Therefore, to use an off-the-shelf projector for monochrome illumination, as desired here, would require removing the spinning filter wheel while keeping the rest of the project intact and then compensating for any artifacts introduced by the conversion from monochrome to color images. Additionally, off-the-shelf projectors are designed to work only at a single frame rate (usually 30 Hz or 60 Hz), and it was likely that I would require much higher frame rates. Moreover, an off-the-shelf projector would restrict our illumination source to the lamp built into the projector, while using a developer's kit DMD would allow us to supply our own laser illumination source of any desired wavelength or power. Consequently the developer's kit proved to be more modular and flexible and better suited to our needs.

Importantly, the developers kit proved to have much lower latency than an off-the-shelf projector. It is well known amongst the computer gaming community that off-the-shelf projectors often have long latencies in excess of 100 ms [106]. Such long latencies would severely the accuracy of the CoLBeRT system, and in fact, such latency did limited the accuracy of an optogenetic illumination system developed by Stirman *et al.*, [147]. That group reports overall latencies of 111 ms to which the projector presumably contributes significantly [148]. Worse, off-the-shelf projectors force the computer to send high-bitdepth full-color images which inherently require higher bandwidth and longer transfer times. CoLBeRT, however, only requires single bit monochrome images, and the DMD developer's kit is able to take advantage of this.

The developer's kit employs fast lossless compression to send images over USB from the computer to a DMD FPGA controller at very high data transfer rates [154]. We estimate that the transfer of one 1024 by 768 pixel monochrome binary image is roughly 5 ms (an effective transfer rate of ~157 Mbit/s). This estimate is based on the time it takes for the DMD software API call to return, so it cannot be ruled out that the latency is in fact larger. For example, the function

call could return before the binary image arrives at the DMD. Nonetheless, 5 ms is consistent with our other measurements. A 5ms transfer time from computer to DMD is consistent with the overall accuracy of the system as measured in the Kaede experiments in Figure 2.5a,b and it is further consistent with the advertised data rate of up to 1.2 Gbit/s provided by the vender of the developer's kit [154][2]. In addition to its simplicity, the developer's kit also offers orders of magnitude improvement in latency.

2.7.3 Schemes for segmentation and targeting

To identify target neurons, I wrote computer vision software that segments the worm based on its outline. However, I also considered schemes to segment the worm based on optical features of the worm's body, such as the location of its eggs or vulva, or based on the fluorescence of individual neurons or markers. Segmenting based on the worm's outline has immediate advantages: under darkfield illumination there is high contrast between the worm and its background which makes locating the worm's outline fairly robust. Additionally, identifying the outline of the worm is computationally simple and is already implemented in many existing computer libraries. Consequently, segmentation based on the worm's outline can be implemented very quickly. In my implementation using OpenCV and Intel's Integrated Performance Primitives, the segmentation completes in under a millisecond.

Segmentation based on outline, however, also has obvious drawbacks. The segmentation fails whenever the worm touches itself, curls on itself or interacts with any object in its environment. Additionally, the high contrast images that can theoretically be attained from darkfield illumination degrade dramatically if the worm is crawling on a substrate that has optical scattering. For this reason much of the

[2]The overall closed-loop latency of the CoLBeRT system could be measured precisely by performing the following experiment inspired by discussions with Marc Gershow: Develop software that instructs CoLBeRT to rapidly generate illumination patterns such that the left side of the illumination image displays a timestamp, e.g. the face of a digital stopwatch, and the right side of the illumination image displays the left region of the most recently acquired video frame. Because the illumination image is projected into the field of view of the camera, every image after the first will show visually the precise time lag between acquisition of one image and the generation of the next. The CoLBeRT system records each of these images. Therefore, to measure the time lag between any two adjacent frames, simply review the recorded images and subtract the timestamp on the right side of the image from the timestamp of the left.

experiments presented here were performed on worms crawling in a viscous liquid which has excellent optical properties. In some experiments, however, it was necessary to image the worm on agar, which acts as a scatterer. In those experiments the segmentation was slightly degraded. Segmentation based on anatomical features or neuronal fluorescence would circumvent these problems.

I concluded that segmentation based on optical features of the worm, such as the location of eggs, the pharynx or the vulva were to too algorithmically complex and computationally intensive. Segmenting based on fluorescence, however, remains appealing. The key challenge posed by segmenting via fluorescence is to acquire enough photons from fluorescing cells to form in image given short exposure times. 1.7 ms exposures were used to capture dark field images of the moving worm. Normal fluorescent imaging, however, requires exposure times of at least 30 ms. Increasing the power of the excitation light or choosing a camera that is more sensitive at low-light may help, but both come with their own tradeoffs. Shining more light on the sample can damage the worm or cause photobleaching, and low-light EMCCD cameras suffer from slower readout times than the faster less sensitive CMOS cameras. As camera technology improves, however, I suspect that segmenting based on fluorescence expression will become a viable option.

2.7.4 Computer vision implementation

Programming Language

In principle, the MindControl computer vision software to segment the worm could have been implemented in any computer programming language. If there were no closed-loop real-time requirements, could have been implemented in a high-level interpretive language such as MATLAB or Python that lends itself to rapid prototyping and that has many pre-existing image processing routines. If latency were less of a priority, LabVIEW would have been a good choice because it lends itself well to graphical user interface (GUI) design, image acquisition and hardware control. The combination, however, of real-time closed-loop processing requirements and the demand for very low-latencies made the programming language C the best option.

C is unparalleled in its speed and in the power it provides to fine-tune all aspects of the code including memory allocation and multithreading. In C I wrote

my own algorithms that performed only the exact computations I necessary, and I was able to optimize the code so as to perform certain computations and memory management tasks immediately prior to acquiring data. Moreover, C worked well with the OpenCV computer vision library [15, 16] which permitted certain operations, such as such gaussian blurs or finding the boundary of the worm, to be hardware accelerated on the central processing unit's (CPU's) onboard vector processing.

C, however, has its own disadvantages. C is a more complex and less forgiving language than most high-level alternatives. It is also ill-suited for writing GUIs and few built-in routines. Moreover, unlike LabVIEW, C is not regularly used amongst scientists for acquiring images or controlling external hardware. This required me to write my own wrapper libraries to interface with our camera and with the DMD controller hardware. Nonetheless, the low-latency requirements superseded other considerations and made C the best choice.

CPUs, GPUs and FPGAs

The MindControl software for CoLBeRT is implemented on an Intel x86 CPU. However, graphical processing units (GPUs) and field programmable gate arrays (FPGAs) are increasingly becoming popular as a way to complement or replace the CPU for high-performance computing applications. GPUs take advantage of dedicated graphics processors to perform vector computations rapidly and in parallel. They are particularly well suited for massively parallelizable computations such as video rendering. Low-latency closed-loop computer vision as required here, however, does not lend itself to massive parallelization. On the contrary, the CoLBeRT system is highly serial– the value of each frame degrades as each millisecond passes, so it only makes sense to process the most recent frame or frames at any given time. As such, the opportunity for GPU acceleration is more limited. Moreover, there are certain disadvantages that come with using a GPU. GPUs can only be interfaced through special programming language frameworks or libraries. Additionally GPUs could potentially increase latency because images have to be loaded from RAM onto a GPU's dedicated memory before the GPU could access it. It remains unclear whether GPUs would provide an overall speed-up over a CPU approach for this application.

Field programmable gate arrays offer another alternative for fast dedicated computations and they have been successfully employed for real-time closed-loop biophysical applications previously[49]. I chose not to explore FPGAs for the following reasons: First, FPGAs require longer development and programming time than CPUs. Second, my software needed to interface with third-party hardware such as the camera and DMD through proprietary APIs. It is unclear whether an FPGA would easily be able to interface with such hardware. FPGAs remain a technology to explore in the future, but for the initial development of CoLBeRT I chose to pursue a CPU-based hardware platform.

Naive or predictive models of worm position

The MindControl software segments worms and sends illumination patterns to the DMD based solely on the most recently acquired image. With few exceptions, the system is entirely "memoryless,"– each frame is segmented independently with no memory of the previous frame. Tossing out such information is inherently a missed opportunity. Moreover, the system is not predictive. There is no model in place to predict where the worm will be 20 ms after an image was acquired.

In the spirit of iterating and rapid prototyping, I chose to pursue a naive segmentation system for this version of the MindControl software. In future versions, however, I hope to implement a predictive model, like the Kalman filter, that takes advantage of the prior state of the worm and uses it to more accurately predict the current position of the worm or even the position of the worm in the near-future. This could dramatically increase accuracy, and it would also allow for sending updated illumination patterns to the DMD device even between acquiring new frames. If such a predictive model proved accurate, it might relax some of the stringent low-latency requirements which would, in turn, open new possibilities for acquiring images and segmenting the worm.

2.7.5 Camera selection

The CoLBeRT system requires a camera that operates at high frame rate and low latency. This restricted our camera options to those cameras that supported a CameraLink interface, as opposed to USB, Firewire or GiGE. Compared to these

other interfaces, the CameraLink interface provides the highest bandwidth and lowest latencies. The CameraLink interface operates direct connects the camera to a dedicated framegrabber that sits in a computer's PCI Express port. The "10-tap Full" CameraLink standard provides a total bandwidth of 5.44 Gbit/s or roughly 680 MB/s. Of the CameraLink cameras, I chose a PhotonFocus CMOS camera, as described in Section 2.4.2. CMOS cameras have a reputation for having faster readout times, but poorer low-light performance compared to EMCCD cameras. Under brightfield illumination low-light conditions was not of concern, while fast readout time was.

2.7.6 Data Storage and organization

For every minute of operation the, CoLBeRT system generates roughly 200MB of video and data to be processed and analyzed at a later date. Additionally, the *Mindcontrol* software is constantly being developed and improved. Keeping track of such large amounts of data from multiple versions of the software was of important consideration. To preserve size, video was downsampled and stored in motion JPEG format, which can be easily imported into MATLAB or played on the open source VLC media player [123].

Data streams were stored in human- and computer-readable YAML formatted text files [44]. The YAML format makes it easy for a human to read out information about each recorded frame such as the time the frame was recorded, the frame number, the state of the CoLBeRT system, the location of the worm's head and tail, the worm's centerline and whether the laser was on or off. YAML is also computer readable and can be imported into MATLAB, C or other programming languages with pre-existing YAML bindings.

Additionally, each data file is automatically stamped with a version hash that uniquely identifies the version of the software that generated it. This allows researchers, after the fact, to reconstruct the state of the software as it was when a given experiment was conducted. This is useful in tracking down bugs and preserving the integrity of the data. The versioning system leverages the open source version control system Git [108]. Code snippets that I wrote to allow C or MATLAB to programatically access the Git version hash are available at `http://stackoverflow.com/a/1843783/200688` and `http://www.`

mathworks.com/matlabcentral/fileexchange/32864-get-git-info , respectively.

2.8 Manuscript Information

2.8.1 Previously Published As

A previous version of this chapter appeared in [92]:

Andrew M Leifer, Christopher Fang-Yen, Marc Gershow, Mark J Alkema, and Aravinthan D T Samuel. Optogenetic manipulation of neural activity in freely moving caenorhabditis elegans. *Nature Methods*, 8(2):147–152, February 2011. ISSN 1548-7105. doi: 10.1038/nmeth.1554. URL http://www.ncbi.nlm.nih.gov/pubmed/21240279. PMID: 21240279

2.8.2 The Author's Contribution

Andrew M. Leifer conducted the experiments described in this chapter and performed all of the analysis, except for the Kaede experiment which was analyzed by Chris Fang-Yen. Andrew M. Leifer designed, built and tested the CoLBeRT system with optics help from Chris Fang-Yen. Andrew M. Leifer conceived of and designed the egg-laying, the habituation and the Kaede experiments, as well as the experiment in Figure 2.1c. Chris Fang-Yen, Mark J. Alkema and Aravinthan D. T. Samuel conceived and designed the other experiments.

The MindControl software and was written by Andrew M. Leifer with mentorship from Marc Gershow. The analysis software was written by Andrew M. Leifer and Chris Fang-Yen. Andrew M. Leifer generated the figures, wrote the captions and wrote the Experimental Setup section, Design Considerations section and portions of the Results and Methods section. The remainder was written by Aravinthan D.T. Samual, Chris Fang-Yen and Andrew M. Leifer.

Motor behaviors are built out of a sequence of movements that evolve through time. From the most basic, such as locomotion, to the most complex, such as playing the piano, the timing of movements is crucial.

Michael A. Long and Michale S. Fee, [109]

3

Bending waves during *Caenorhabditis elegans* locomotion are driven and organized by proprioceptive coupling

3.1 Introduction

How neural circuits give rise to coordinated rhythmic behaviors such as locomotion remains a fundamental question in systems neuroscience [37]. Classic studies sought the neuromuscular basis of locomotion in aquatic swimmers such as lamprey and leech [31, 42, 53, 89, 113]. In these systems, the concept of a Central Pattern Generator (CPG) is commonly evoked to explain rhythmic behavior [37, 113]. The CPG hypothesis is supported by observations that motor neurons in each body segment continue to exhibit rhythmic activity even after pruning all inputs [31, 89, 125].

Whereas a CPG could produce rhythmic behavior, motor circuits still need to respond to sensory inputs to deliver precise and flexible control of body move-

ment [37]. In leech, muscle activity can be coordinated among segments by sensory feedback even after cutting neuronal couplings between segments [172]. In *Drosophila* larva, specific classes of mechanosensory neurons that tile the body contribute to organizing peristaltic waves during locomotion [26, 75, 142]. In *C. elegans*, the shape and speed of bending waves adapt to the mechanical load imposed by the environment [10, 45], and mutations in a mechanosensitive channel (TRP-4) acting in the DVA interneuron perturb the amplitude and frequency of body undulation [94].

Here, we sought a biophysical characterization of the role of proprioceptive feedback in the *C. elegans* locomotory circuit by combining microfluidics and optical neurophysiology [28, 30, 98, 107, 174]. We discovered that stretch-sensitive coupling between adjacent body regions represents the key mechanism for driving bending waves along the worm body during forward locomotion.

3.2 Results

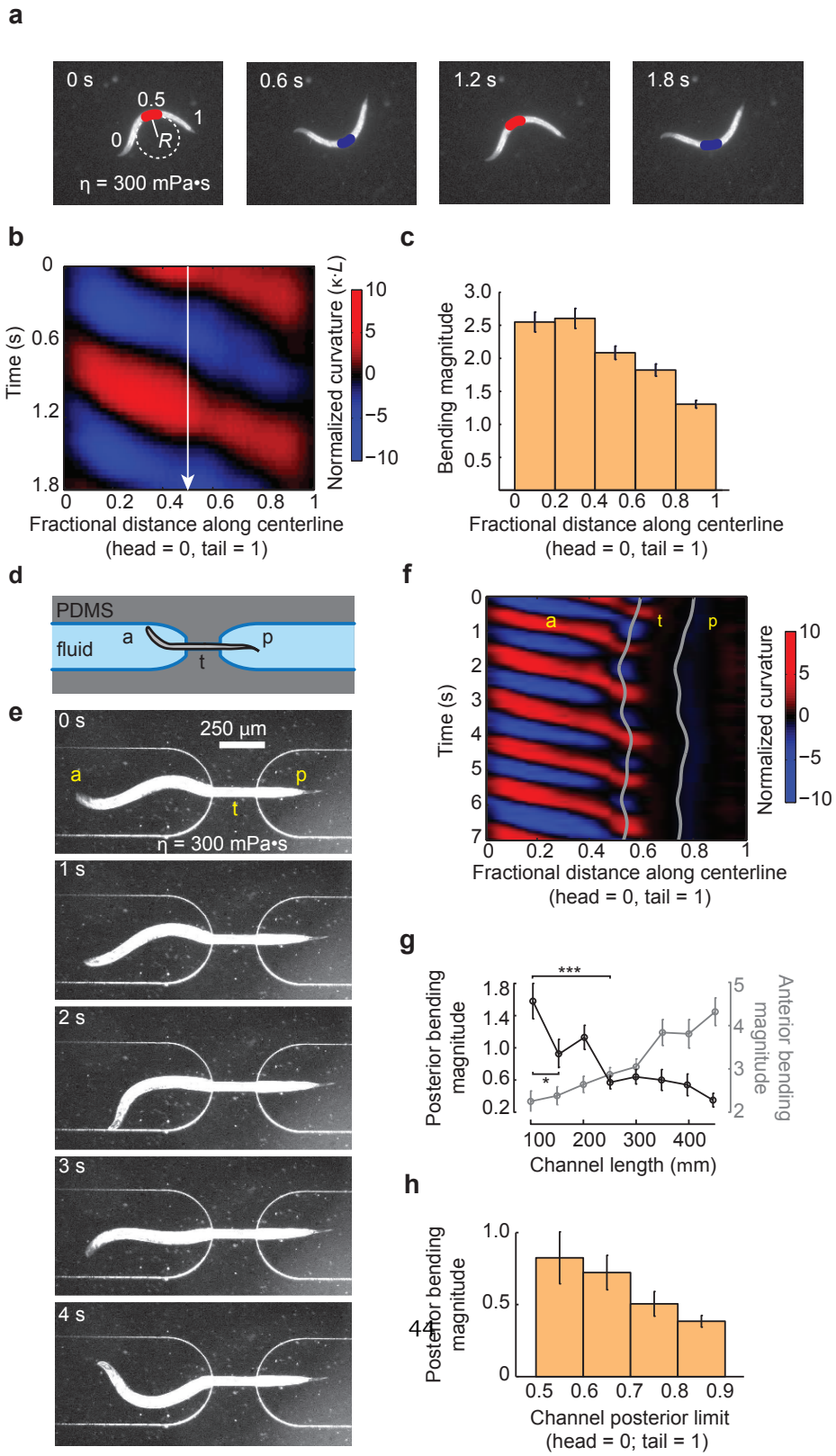
3.2.1 The bending of one body region requires the bending of its anterior neighbor

C. elegans moves forward by propagating dorsal-ventral body bending waves from head to tail. The detailed kinematics of bending waves can be quantified using the time-varying curvature measured at each point along the body centerline over time (Fig. 3.1a). To compare data from different worms, we normalized distance along the centerline by measuring fractional distance from head to tail (head $=$ 0; tail $=$ 1). Each body region alternates between positive (red) and negative (blue) curvature, and bands of curvature propagate from head to tail as shown in a kymograph (Fig. 3.1b).

First, we sought to determine whether the motor activity in one body region depends on the bending of neighboring body regions. To test this, we designed microfluidic devices that enabled us to immobilize body regions of varying length along the middle of a young adult worm (Fig. 3.1d-e and Video 3.5.1). Our first device trapped the center of a worm in a narrow straight channel. The region of the worm's body inside the channel was restrained, while regions of the worm's

Figure 3.1 ***(following page)*:** Bending of posterior regions requires anterior bending. (**a**) Video images of a worm swimming forward. Bending is quantified by measuring the radius of curvature R at each point along the centerline. Position along the centerline is measured in normalized coordinates using the fractional distance from head to tail (head = 0; tail = 1). A red-blue colormap illustrates alternating curvatures at fractional distance = 0.5. (**b**) Kymograph of time-varying curvature illustrating retrograde bending waves along the worm represented in non-dimensional units. To do this, the curvature at each point along the centerline, $\kappa = 1/R$, is multiplied by worm length, L. (**c**) Bending magnitude along the body of a wide-type free swimming worm, measured as the standard deviation of normalized curvature over time. $n = 18$ worms, mean $\pm$ one standard error. (**d**) Schematic of microfluidic device. a stands for anterior region, p stands for posterior region, and t stands for trapped region of a worm. (**e**) Video images of a wide-type young adult worm exhibiting forward undulatory gait inside the microfluidic device (also see Video 3.5.1). The channel divides the worm body into unrestrained anterior, posterior, and trapped middle regions. (**f**) Kymograph of time-varying curvature along the body of the worm shown in (**e**). Gray lines mark the anterior and posterior limits of the straight channel. (**g**) Bending magnitude of a posterior and an anterior body region (~0.15 worm length) adjacent to the channel, measured as the standard deviation of time-varying normalized curvature, is plotted as a function of the length of the trapped region. $n \geq 10$ worms for each condition, mean 15 $\pm$ one standard error. Position of the posterior limit of the channel is 0.7 $\pm$ 0.1 (mean $\pm$ standard deviation) for each condition, measured as the fractional distance from head to tail. *$P < 0.05$, ***$P = 0.0001$, Mann-Whitney U test. (**h**) Bending magnitude of a posterior body region (mean $\pm$ one standard error) as a function of the position of the posterior limit of the channel. We measured 64 bouts of forward movement trapped in different channel positions from 20 worms. Channel length is 300μm.

Figure 3.1: (continued)



body either anterior or posterior to the channel were free to bend (Fig. 3.1d-e). We used a channel diameter (40μm) that was sufficient to immobilize the trapped region (worm diameter is $54 \pm 4\mu$m; mean $\pm$ SD) with minimum constriction.

We consistently recorded bouts of forward movement ($>$ 10 s) when we set the posterior limit of the channel at 0.7 ± 0.1 in fractional worm length along the body. Bending waves would propagate normally to the anterior limit of the channel (gray data points in Fig. 3.1g). Short channels (100μm long) did not affect wave propagation along the worm body; the bending wave that emerged from the posterior limit of the channel (black data points in Fig. 3.1g) exhibited similar amplitude as a free swimming worm (Fig. 3.1c). However, increasing channel length beyond 200μm significantly diminished the bending amplitude in the posterior body region (Fig. 3.1e-g). Fixing the channel length, but moving it toward the tail also reduced the posterior bending amplitude (Fig. 3.1h, $R = -0.24$, $p < 0.05$, Spearman's rank correlation test).

To determine whether immobilization directly affects muscle activity in body regions within and posterior to the channel, we quantified intracellular calcium dynamics in the muscle cells of transgenic worms (P*myo3::G-CaMP3::RFP*) expressing the calcium indicator G-CaMP3 [153] and RFP in all body wall muscles (Supplementary Fig. 3.6 and Video 3.5.2). While muscle cells anterior to the channel exhibited strong rhythmic intracellular calcium dynamics during the propagation of bending waves, muscle cells within and posterior to the channel had much lower levels of calcium dynamics (Supplementary Fig. 3.6).

Taken together, these results suggest that immobilizing a body region lowers motor activity within and posterior to that region, thereby disrupting bending wave propagation. Motor activity in one body region seems to require active bending of anterior regions extending ~200 μm.

3.2.2 Muscle activity is positively correlated with the curvature of adjacent anterior neighbors

To further explore how the bending of adjacent body regions is coupled, we designed microfluidic devices that trapped the middle region of a worm at defined curvatures (Fig. 3.2a,c). Here, we used channels that were at least 250μm long to prevent bending waves from propagating into the unrestrained posterior part. The

unrestrained posterior region exhibited static curvature in the same direction as that imposed on the middle region trapped by the channel (Fig. 3.2a-b and Video 3.5.3). By using channels with different curvatures, we found that the curvature of the posterior region increased linearly with the imposed curvature on the trapped middle region with slope $0.62 \pm 0.03L$ (Fig. 3.2d).

We sought to verify that the static curvature of the posterior unrestrained region was driven by muscle activity, and not through passive mechanical properties of the worm body. First, we used transgenic worms (*Pmyo3::NpHR*) that express halorhodopsin [64] in all body wall muscles. When we induced muscle relaxation in the unrestrained posterior region with green light, we found that the tail reversibly straightened during illumination (Fig. 3.2e-g and Video 3.5.4). Second, we directly monitored muscle activity in the curved posterior region using 5 transgenic worms (P*myo3::G-CaMP3::RFP*) that expressed both G-CaMP3 and RFP in all body wall muscle cells (Fig. 3.2h). Using this strain, intracellular calcium levels within a cell can be inferred from the ratio of green to red fluorescence–the higher the ratio, the higher the intracellular calcium concentration. In the posterior region emerging from the channel, we consistently measured higher calcium levels in the muscle cells on the inner side than the outer side of the curved body (Fig. 3.2i-k). Third, when the whole animal was paralyzed with sodium azide, the body regions emerging from the curved channel remained straight (Video 3.5.6). These experiments suggest that the static curvature of the posterior unrestrained region is due to a fixed pattern of motor circuit activity.

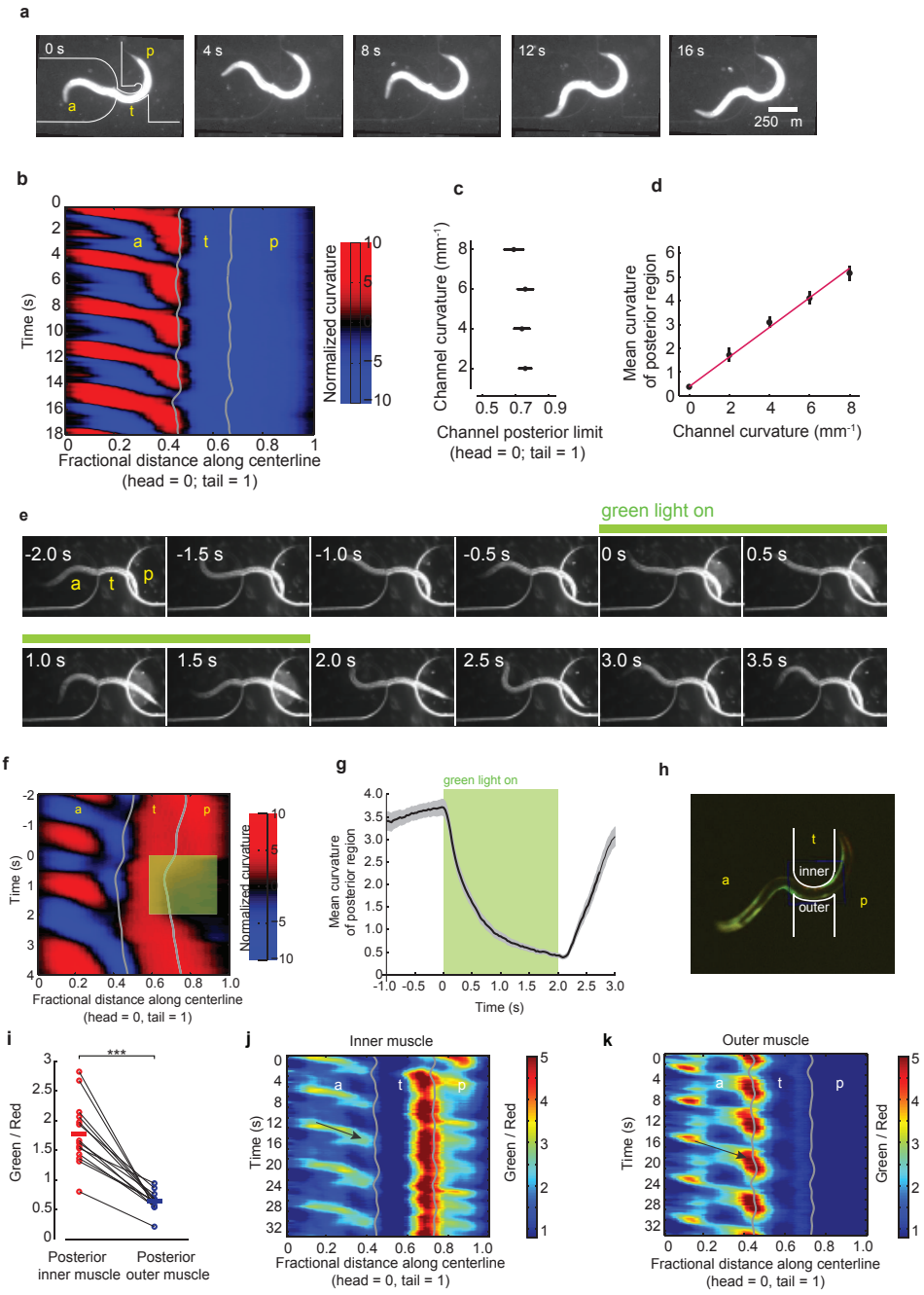
Taken together, our results suggest that proprioceptive coupling contributes to propagating the bending signal along the worm body during forward movement. Through positive stretch-sensitive feedback, posterior regions are compelled to bend in the same direction and in proportion to the bending of adjacent anterior regions.

3.2.3 Post-channel body curvature follows channel curvature with a viscosity- dependent delay

To characterize the temporal dynamics of proprioceptive coupling, we measured the time lag between the bending in one body region and the induced bending in

Figure 3.2 ***(following page)*****:** Bending of posterior regions is positively correlated with anterior bending. (**a**) Video images of a worm exhibiting forward undulatory gait while partially constrained in a curved microfluidic channel (also see Video 3.5.3). (**b**) Kymograph of normalized curvature of the worm shown in (**a**). Gray lines show anterior and posterior limits of the curved channel. (**c**) Positions of the posterior limit of the curved channel (mean $\pm$ standard deviation). $n \geq 8$ worms for each condition, (**d**) The curvature of the unrestrained posterior body region, measured as a spatial average from the posterior limit of the channel to the tail and a temporal average over bouts of forward movement, is plotted as a function of channel curvature. Each data point (mean $\pm$ one standard error) represents data from at least 8 animals. Magenta line is the linear least square fit of the data with a slope 0.62 ± 0.03 L. (**e**) Video images of a transgenic worm (P*myo3::NpHR*) partially constrained in a curved microfluidic channel. The green bar indicates a 2 s interval during which the posterior body region emerging from the channel was illuminated by green light (also see Video 3.5.4). (**f**) Kymograph of normalized curvature of the animal shown in (**e**). Green shading indicates the body region and duration of green light illumination. (**g**) Mean curvature $\pm$ one standard error of the posterior region emerging from the curved channels as shown in (**a**) during green light illumination (~30 measurements using 6 worms). (**h**) Calcium imaging of body wall muscles in a partially constrained transgenic worm (P*myo3::G-CaMP3::RFP*) in a curved channel. Red fluorescence from RFP constitutes the reference signal. Green fluorescence from G-CaMP3 indicates intracellular calcium levels. The contours of the microfluidic channel are drawn in white (also see Video 3.5.5). (**i**) Comparison of the ratio of green fluorescence to red fluorescence intensity emitted from inner and outer muscles of the posterior body region. Each data point represents a spatial average of the ratio over a posterior body region (~0.2 worm length) adjacent to the channel and a temporal average over a bout of forward movement. Solid lines indicate population mean. Among 14 measurements from six worms, six measurements restrict dorsal muscles on the inner side. ***$P = 0.00001$, Mann-Whitney U test. (**j,k**) Representative ratiometric kymograph of calcium levels in inner (**j**) and outer (**k**) muscle cells of a worm trapped in the device shown in (**h**). Higher/lower ratios of green fluorescence to red fluorescence in each set of body wall muscles indicate higher/lower intracellular calcium levels. Arrows highlight one calcium wave that propagates from the head to the anterior limit of the curved channel along the inner musculature (**j**) and outer musculature (**k**).

Figure 3.2: (continued)



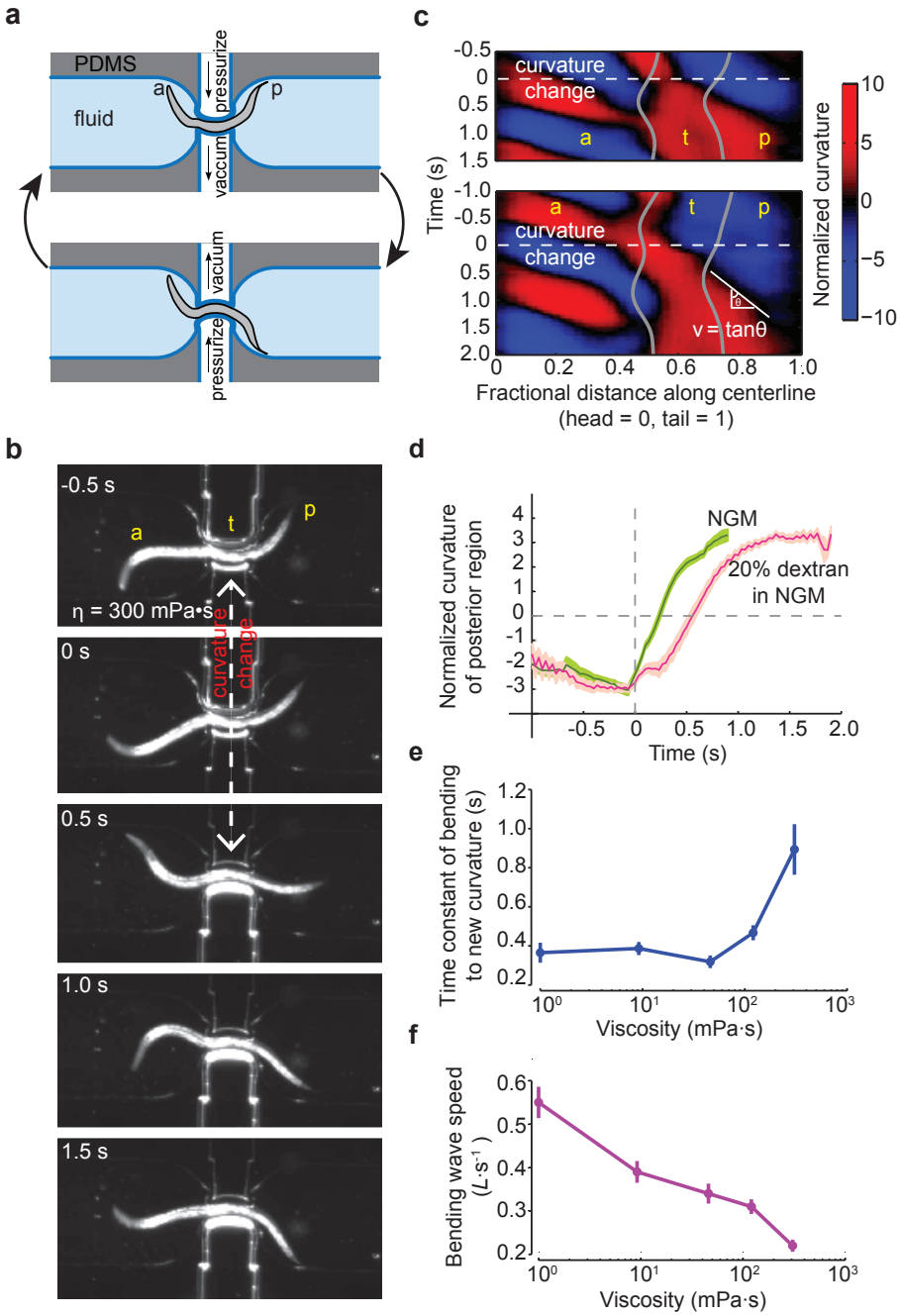
the neighboring posterior region. To do this, we designed pneumatic microfluidic devices to rapidly change the curvature of one region of a worm. We flanked both sides of a thin channel with two independently controllable inflatable chambers (Fig. 3.3a). By simultaneously pressurizing one chamber while depressurizing the other, we were able to induce rapid curvature changes in a specific region of a trapped worm.

As with the static curved channels, we found that the curvature of the worm body posterior to the dynamic channel was positively correlated with channel curvature. Switching channel curvature to either side induced a switch in the curvature of the posterior body region (Fig. 3.3b-c and Video 3.5.7). This result also underscores dorsal-ventral symmetry in the coupling mechanism between adjacent body regions.

We found that the switch in curvature of the posterior unrestrained region propagated with measurable speed from the posterior limit of the channel to the tail, consistent with the flow of a retrograde bending signal (Fig. 3.3c-f). We sought to determine whether the delayed bending of the posterior region represented mechanical damping by the external viscous fluid and/or internal delays within the neuromuscular network. To do this, we studied worms that were immersed in fluids of different viscosity (Fig. 3.3d-f). We found that the bending delay was roughly constant, ~300 ms, in fluids ranging from 1 mPa·s (the viscosity of water) to ~100 mPa·s. In more viscous fluids, the bending delay began to increase, becoming ~1 s at 300 mPa·s. These results suggest that ~300 ms represents an upper bound for delays within the neuromuscular network, which become rate-limiting at low viscosities. Interestingly, ~300 ms also coincides with the undulation period for *C. elegans* swimming in water, and may represent the time constant that sets the upper limit to undulation frequency. Delays within the neuromuscular network might reflect signaling delays in synaptic transmission and/or the limiting speed of muscle contraction.

Figure 3.3 ***(following page)*****:** Pneumatic microfluidic device for manipulating body curvature. (**a**) Schematic of the pneumatic microfluidic device. The channel is flanked by two chambers. By alternatively pressurizing one chamber while depressurizing the other, the curvature of a region of a trapped worm is rapidly switched. (**b**) Video images of a partially immobilized wild-type worm. At $t = 0$ s, the channel starts to change its curvature (also see Video 3.5.7). (**c**) Two representative curvature kymographs of a worm trapped in the pneumatic channel. Gray lines mark the anterior and posterior limits of the curved channels. White dashed lines at $t = 0$ s mark the induced change in channel curvature from negative (color blue) to positive (color red). While the unrestricted anterior body region exhibits opposite bending activities in the two kymographs, this difference did not affect the dynamics of the induced curvature change in the unrestricted posterior body region. The bending wave that shifts the posterior region from negative to positive curvature propagates with velocity equal to the slope of the zero crossing in curvature (color black) as shown.
(**d**) The time course of curvature change in the immediate posterior region (~0.1 worm length) emerging from the pneumatic channel after the switch of channel curvature at $t = 0$ s. The two curves correspond to experiments conducted in two different viscosities (NGM buffer and 20% dextran in NGM). Error bars indicate one standard error. (**e**) The time constant for relaxation of the posterior region to new curvatures obtained by fitting exponentials to time courses as shown in (**d**). Each data point represents at least 30 measurements from five worms. Error bars indicate 95% confidence interval to the exponential fits. (**f**) The speed of the bending wave following induced changes in channel curvature as a function of fluid viscosity. Error bars indicate one standard error.

Figure 3.3: (continued)



3.2.4 Stretch-sensitive feedback requires cholinergic motor neurons

Cholinergic motor neurons that innervate the ventral and dorsal muscle cells are required for *C. elegans* locomotion [22, 92]. B-type cholinergic motor neurons are required for forward locomotion, whereas A-type cholinergic motor neurons are required for backward movement [22]. We asked whether the cholinergic neurons play a role in the proprioceptive coupling that propagates the bending signal along the worm body. First, we trapped transgenic worms (P*unc17::NpHR*) that expressed halorhodopsin in all cholinergic motor neurons in the pneumatic microfluidic devices and illuminated them with green light. We found that deactivating the cholinergic neurons prevented posterior body regions from following induced changes in the curvature of the anterior region (Fig. 3.4 and Video 3.5.8). Instead, optogenetic inactivation of the cholinergic neurons arrests the worm in the posture immediately preceding green light illumination.

Next, we studied *vab-7* mutants, whose dorsal B-type cholinergic motor neurons (DB) reverse the direction of their processes [43]. During unrestrained forward movement, the bending wave in the anterior of *vab-7* mutants exhibits both dorsal and ventral curvatures, whereas the bending wave that propagates to posterior regions exhibits only ventral curvatures (Supplementary Fig. 3.7a, c-d and Video 3.5.9). In other words, only ventral bending waves propagate through the worm body. When we trapped *vab-7* mutants in the pneumatic microfluidic device, we found that the posterior region of the worm was unable to follow the bending of the channel to the dorsal side (Supplementary Fig. 3.7b, e-f and Video 3.5.10). Taken together, these results suggest that the B-type cholinergic motor neurons are required to close a proprioceptive feedback loop that drives the bending signal along the motor circuit during forward movement.

3.2.5 Body wall muscles exhibit hysteresis

Deactivating cholinergic motor neurons in transgenic worms (P*unc17::NpHR*) locked the worm in whatever bending posture the worm had adopted immediately preceding illumination (Fig. 3.4 and [92]). Accordingly, we ask whether

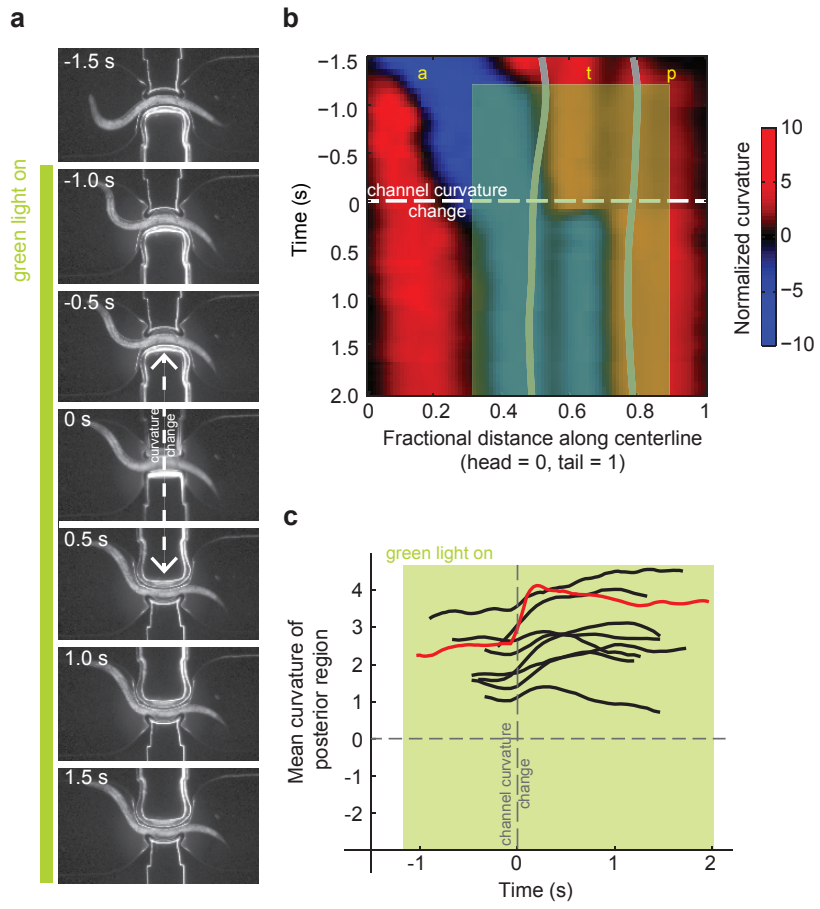


Figure 3.4: Optogenetic inactivation of cholinergic motor neurons. (**a**) Video images of a transgenic worm (P*unc17::NpHR*) partially trapped in a pneumatic microfluidic channel. Green bar indicates the duration of green light illumination of the middle portion of the worm before and after induced change in channel curvature at t = 0 s. As a result, the curvature of the tail failed to follow the curvature change of the channel (also see Video 3.5.8). (**b**) Curvature kymograph of the transgenic worm trapped in the channel as shown in (**a**). Green shading indicates the body region and duration of green light illumination. (**c**) Curvature of the posterior body region, measured as an average from the posterior limit of the channel to the tail, during onset of illumination (green shading) and the induced change in curvature of the middle region at t = 0 (dashed line). Representative data from five worms were shown. Red curve corresponds to the experiment shown in (**a**) and (**b**).

C. elegans muscles can sustain contraction even in the absence of motor neuron inputs.

To test this possibility, we optogenetically stimulated body segments in transgenic worms (P*myo- 3::ChR2*) expressing Channelrhodopsin-2 in body wall muscles while eliminating motor neuron inputs. To remove motor neuron inputs, we treated transgenic worms with ivermectin, which hyperpolarizes the motor circuit by activating glutamate gated chloride channels [34, 38], but does not directly affect muscle cells [66]. When we optogenetically induced body bending in paralyzed worms, the bend would persist long after turning off the illumination (Supplementary Fig. 3.8a-b and Video 3.5.11). The bend would gradually relax over ~40 s, but often in a series of abrupt jumps (Supplementary Fig. 3.8c). We observed similar a phenomenon when ivermectin treated worms are in the *unc-13(s69)* background (Video 3.5.12), a loss of function mutation that prevents synaptic transmission from both GABAergic and cholinergic motor neurons to muscles [132]. Taken together, these results suggest that *C. elegans* body wall muscles indeed exhibit a form of hysteresis: they can maintain stable levels of contraction long after stimulation.

3.3 Discussion

Russell and Byerly noted that the cholinergic motor neurons have long and synapse-free processes that extend along the ventral and dorsal nerve cords. They speculated that these processes might represent stretch-sensitive antennae to detect changes in body posture (cited in [24, 161] and Supplementary Fig. 3.5b). In theoretical models of the worm motor circuit, Niebur and Erdos [121] used proprioceptive coupling hypotheses as a mechanism to propagate bending signals. Here, we have shown that cholinergic motor neurons are required to transduce stretch-sensitive signals, and that this form of proprioceptive coupling represents a key mechanism for propagating bending signals along the *C. elegans* body during forward locomotion. Posterior body segments are compelled to bend in the same direction as anterior segments through stretch-sensitive signals transduced by cholinergic motor neurons. Proprioceptive coupling could explain the organization of the undulatory gait without the need to invoke ensembles of center pattern generators (CPG) along the motor circuit, like body segments in lamprey

and leech are thought to behave [42].

The small size and experimental accessibility of the *C. elegans* motor circuit suggest that it might be possible to build computational models of locomotion that integrate the dynamics of all neuronal and muscle components. Our results suggest that any computational framework for the *C. elegans* locomotory circuit must integrate the biomechanics of undulatory movement itself with neuromuscular activity. In *C. elegans*, the motor circuit organizes the undulatory gait for forward locomotion by both detecting and driving bending activity.

3.4 Material and Methods

3.4.1 Worm strains

Wild-type worms (N2 Bristol) were cultivated at 20 °C using standard methods. We performed all experiments using young adult worms within a few hours after their final molt. The *vab-7* mutant strain was obtained from the *C. elegans* Genetics Center (Minneapolis, MN, US). The *unc-13(s69)* mutant strain was a gift from E. Jorgensen.

Transgenic worms carrying extrachromosomal arrays (P*myo3::G-CaMP3:RFP*) were used for calcium imaging in body wall muscle activities. The promoter sequence of the P*myo-3* gene [91] and the coding sequences of *GCaMP3* and tagRFP-T were ligated to the pSM backbone. The *G-CaMP3* and *tagRFP-T* sequences were separated by a SL2 trans-splicing site. The construct was injected into N2 worms to obtain extrachromosomal arrays.

The transgenic worms used in all optogenetic experiments were cultivated in the dark at 20 °C on NGM plates with Escherichia coli OP50 and all-trans retinal. We made OP50-retinal plates by seeding each 6-cm NGM plate with 250 μl OP50 and 1 μl 100 mM retinal in ethanol. Strains used are ZX444 [*lin-15*(n765ts); *zxEx29* (P*myo-3::NpHR::ECFP*; *lin-15*+)] , ZM 5016 *hpls178* [P*unc-17::NpHR::ECFP*] and ZM5398 *hpls199*[*Pmyo3::ChR2::EGFP*] . ZX444 was a gift from A. Gottschalk and the extrachromosomal array was integrated by cobalt-60 irradiation and outcrossing three times to wild-type worm. Strains ZM 5016 and ZM5398 were gifts from M. Zhen. The strain *unc-13*(s69); *hpls199*[P*myo3::ChR2::EGFP*] was made by crossing *unc-13*(s69) with *hpls199*.

3.4.2 Microfluidic devices

Microfluidic devices were fabricated using standard soft lithography. Each design was drawn in Clewin and sent to a laser-printing service (CAD/Art Services, Inc. Bandon, OR). A master was created by patterning features of SU-8 negative photoresist (Microchem Corp., Newton, MA, US) on a silicon wafer using photolithography. This master was then used to mold microfluidic channels in PDMS. To facilitate the release of the PDMS device from the master, we treated the master with vapor of tridecafluoro(1,1,2,2 tetrahydroooctyl) trichlorosilane (Gelest, Inc., Philadelphia, PA, US) inside a vacuum chamber. The PDMS prepolymer was mixed with Sylgard 184, its curing agent, at specific weight ratios (20:1 for the 10 pneumatic microfluidic device and 10:1 for other devices). After pouring the PDMS prepolymer over the master, we cured the PDMS at 60 °C for 8 h and peeled the PDMS slab from the master. A circular biopsy punch (1.5 mm in diameter, Shoney Scientific Inc., Waukesha, WI, US) was used to create inlets and outlets for the microfluidic channels. The PDMS slab was bonded to a glass coverslip by treating the surfaces of the glass and the PDMS slab with air plasma for two minutes and 30 s respectively. The PDMS was bonded to another PDMS slab in the pneumatic microfludic device.

We loaded the microfluidic channel with NGM buffer or dextran solution [~20 channel and worm position within each channel was manually controlled by syringes connected to polyethylene tubing (1.57 mm outer diameter and 1.22 mm inner diameter). In the calcium imaging experiments, channel depth was ~38 μm. In all other cases, channel depth was ~85 μm. In the pneumatic microfluidic device, the channel was flanked by two chambers that could be alternatively pressurized and depressurized with a valve system under computer control using custom software written in LabVIEW (National Instruments, Austin, TX, US).

3.4.3 Measuring locomotion of partially immobilized worms

Experiments were performed on Nikon microscopes (TE2000 or Eclipse LV150) under 4X magnification with dark field illumination. Image sequences were taken by a CCD camera (Imaging Source) and recorded on a computer at 30 Hz using IC Capture software (Imaging Source).

Image analysis was performed using custom software written in MATLAB

(MathWorks, Inc. Natick, MA, US) following methods described in [45]. Briefly, we identified image sequences in which the worm persistently exhibited its forward undulatory gait, i.e., bending waves propagated backward from the head. After background subtraction to eliminate features of the microfluidic channel, we filtered and thresholded each image to obtain a binary image. The head and tail were identified as the points of maximum convex curvature on the worm boundary. A centerline extending from the head to the tail of the worm was calculated so that each point is roughly equidistant to nearest boundary points on two sides of the animal. The centerline was then fit by a least-squares cubic smoothing spline. Curvature, by definition, was calculated as the magnitude of the derivative of the unit vector tangent to the centerline with respect to the body coordinate along the centerline. In the experiment using pneumatic microfludic devices, the speed of the bending wave of the body region posterior to the channel was calculated using least-square linear fits to the zero crossings of curvature (Fig. 3.4c).

3.4.4 Ivermectin-induced paralysis.

Young adult transgenic worms (*hpls199* [*Pmyo3::ChR2::EGFP*] or *unc-13*; *hpls199*) were placed in ivermectin solution (0.02 mg/mL in NGM) sandwiched between two glass sides separated by 127 μm. Worms were free to swim until completely paralyzed after ~30 min, when they were used for optogenetic stimulation.

3.4.5 Calcium imaging of body wall muscle activities.

GCaMP3 and RFP were excited by LEDs filtered at 448-492nm and 554-572nm respectively using Semrock single-bandpass filters. Fluorescence emission was recorded through an Olympus MVX Plan Apochromat 2X objective (working distance 20mm, numerical aperture 0.5). The fluorescence image was split by a Cairns Optosplit II Image Splitter and the two images (green channel, 499-525 nm; red channel, 581- 619 nm) were projected onto two halves of an Andor iXon 885 EM-CCD camera. A DinoLite Pro AM413T USB camera was used to track the worm using Worm Tracker 2.0 software developed by the Schafer lab. Zaber T-LSR075A Motorized Linear Slides give automated x-y stage movement. Imaging sequences were recorded on a computer at 10 Hz using Andor Solis software and converted into TIFF files using ImageJ. Images were then analyzed using custom-written

MATLAB scripts. Briefly, the two split images were re-aligned and the calcium activities of muscles were calculated as the ratio of green to red fluorescence emission intensities. The true emission intensities from the two channels were calculated using the following formulas: True green = green measured – green background; True red = red measured – red background – 0.153*True green. There is 15.3% bleedthrough from the green to the red channel.

3.4.6 Optogenetic stimulation.

We used two optical setups to stimulate transgenic worms expressing Channelrhodopsin (ChR2) or Halorhodopsin (Halo). Experiments with the pneumatic microfluidic device were conducted on a Nikon microscope (Eclipse LV150) under 10X magnification with dark field illumination. A mercury arc lamp with green filter and field 12 diaphragm was used to illuminate the worm with controlled spot size. Rhodamine in the microfluidic channel (10 μM) allowed us to directly visualize the area and duration of green light illumination. Other optogenetic experiments were conducted on a modified version of the CoLBeRT system, described in [92]. Briefly, the CoLBeRT system consists of an inverted Nikon microscope (TE2000), blue and green diode pumped solid state lasers, a high speed CCD camera, and a digital micromirror device all under the control of the open source MindControl software. Each worm was imaged under red light with dark field illumination, and the digital micromirror device reflected laser light to shine on targeted cells or regions of the worm. For the *unc-13*; P*myo3::ChR2* and ivermectin-treated worm experiments, the CoLBeRT system was modified to programmatically control laser intensity. To do this, the MindControl software interfaces with custom-written LabVIEW software that modulates the analog voltage signal sent to the laser power controllers via a LabJack U3-HV digital-to-analog converter. MindControl and associated software is available at `http://github.com/samuellab`.

3.5 Accompanying Video

3.5.1 Video 1

A young adult N2 worm was swimming forward in dextran solution [20% dextran in NGM (wt/vol)] while the middle region of its body was constrained in

a straight microfluidic channel. Body undulation can propagate to the anterior limit of the channel, but bending waves emerging from the posterior limit of the channel were hardly detected and the unrestricted posterior body region remained straight (QuickTime; 12.2 MB).

3.5.2 Video 2

A young adult transgenic worm expressing calcium indicators in body wall muscle cells (P*myo3::G-CaMP3::RFP*) was swimming forward in dextran solution [15% dextran in NGM] while the middle region of its body was constrained in a straight microfluidic channel. White lines show the boundary of the channel. Pseudo-color shows the ratio of green fluorescence (emitted from G-CaMP3) to red fluorescence (from RFP). Higher ratio, which indicates higher calcium activities, is represented by a larger value of the red component in the muscle color map. Muscle cells within and posterior to the channel had lower levels of intracellular calcium dynamics than muscle cells anterior to the channel that drove bending waves (QuickTime; 0.56 MB).

3.5.3 Video 3

A young adult N2 worm was swimming forward in dextran solution (20% dextran in NGM) while the middle region of its body was constrained in a curved microfluidic channel. The unrestricted posterior region of the worm exhibited static curvature in the same direction as that imposed on the middle region by the channel (QuickTime; 0.82 MB).

3.5.4 Video 4

A young adult transgenic worm expressing halorhodopsin in its body wall muscles (Pmyo3::NpHR) was swimming forward in dextran solution (20% dextran in NGM) while the middle region of its body was constrained in a curved microfluidic channel. A bright circular spot appeared when the posterior body region was illuminated by green light. During illumination, the unstrained posterior body region was reversibly straightened (QuickTime; 1.6 MB).

3.5.5 Video 5

A young adult transgenic worm that expresses calcium indicators in body wall muscle cells (P*myo3::G-CaMP3::RFP*) was swimming forward in dextran solution (15% dextran in NGM) while the middle region of its body was constrained in a curved microfluidic channel. White lines show the boundary of the channel. Pseudo-color shows the ratio of green fluorescence (from G-CaMP3) to red fluorescence (from RFP). Higher ratio is represented by a larger value of the red component in the muscle color map. During forward locomotion, the muscle cells at the inner side of the posterior body region emerging from the channel consistently exhibited higher calcium activities than the outer side (QuickTime; 5.5 MB).

3.5.6 Video 6

After being fully paralyzed by sodium azide, a young adult N2 worm was partially constrained in a curved pneumatic microfluidic channel. The unrestrained body regions emerged from the channel always remained straight (QuickTime; 0.46 MB).

3.5.7 Video 7

Young adult N2 worms were moving forward while the middle regions of their bodies were constrained in pneumatic microfluidic channels. In the first half of the video, a worm was swimming forward in NGM buffer solution. Switching the bending direction of the channel quickly induced the switch in the bending direction of the posterior body region. In the second half of the video, a worm was swimming forward in more viscous fluid (20% dextran solution in NGM). In this case, there was a more significant delay between switching the bending of the channel and switching the bending direction in the posterior body region (QuickTime; 5.7 MB).

3.5.8 Video 8

A young adult transgenic worm (P*unc17::NpHR*) that expressed halorhodopsin in all cholinergic motor neurons was moving forward while the middle region of the body was constrained in an pneumatic microfluidic channel. During green light

illumination, the unrestrained posterior body region failed to follow the switch of channel curvature (QuickTime; 1.7 MB).

3.5.9 Video 9

A young adult *vab-7* mutant was moving freely in 20% dextran solution. The ventral and dorsal side of the worm can be identified by the location of eggs from the dark field illumination. During forward locomotion, dorsal bending waves could not propagate throughout the body and the posterior body region always curved ventrally (QuickTime; 1.3 MB).

3.5.10 Video 10

A young adult *vab-7* mutant was moving forward while the middle region of its body was constrained in an inflatable microfluidic channel. Bending the channel to the dorsal side failed to induce the switch in the bending direction of the posterior body region (QuickTime; 1.2 MB).

3.5.11 Video 11

This video shows 2 s blue light illumination of an ivermectin-treated paralyzed transgenic worm (P*myo3::ChR2*) that expressed Channelrhodopsin-2 in body wall muscles. When DLP was on, blue light illuminated a rectangular region near the middle of the worm and induced body bending. When DLP was off, there was no light stimulation but the bending persisted for a long period (QuickTime; 3.8 MB).

3.5.12 Video 12

This video shows 2 s blue light illumination of an ivermectin-treated paralyzed transgenic worm (P*myo3::ChR2*) in the *unc-13*(s69) background. When DLP was on, blue light illuminated a rectangular region near the middle of the worm and induced body bending. When DLP was off, there was no light stimulation but the bending persisted for a long period.

3.6 Supplementary Materials

3.6.1 Parallels to songbird motor pathway

The work presented here tries to address previously open questions about how the sequential activation of muscles in *C. elegans* is coordinated and controlled and what role muscles, motorneurons and feedback play, respectively. Prior to this work, existing data was insufficient to discriminate between conflicting models. For example, in one model sequential activation of motorneurons could drive muscle contraction without any proprioceptive feedback. In another model, muscles could propagate bending waves by sequentially activating neighboring muscles via gap junctions without any need for proprioceptive feedback and with motorneurons playing only a modulatory role. Yet other models suggested that proprioceptive feedback was the driving force behind wave propagation, which is in fact consistent with the observations shown here.

By selectively inhibiting regions of muscles, motorneurons and physical restraining different regions of the worms body, this work rules supports a model based on proprioceptive feedback. Namely, motorneurons are thought to be activated by yet-to-be discovered stretch sensors which in turn drive muscles in one region to contract, which in turn activate stretch sensors, which finally induce posterior motorneurons to contract— perpetuating a cycle that propagates bending waves.

The questions posed and the experimental approach taken in this work has parallels to prior work done by Michael Long and Michale Fee studying birdsong in zebra finch [109]. In that work, Long and Fee were interested in similar types of questions related to songbird motor pathway. They asked what role interneurons in a brain region HVC and motorneurons in a region RA played, respectively, in sequential muscle activation required for birdsong. Prior to their work, it was hypothesized that HVC sequentially initiated small subsequences of shorter timescale motor activation that were handled entirely within RA. By selectively cooling either RA or HVC and observing slowing in the resulting birdsong they show that HVC directly induces each motorneuron activation in RA, even on short time scales. They also show that two independent hemispheres of HVC take turns controlling the motor sequence and they provide evidence that feedback downstream

of RA is used to keep the two hemispheres synchronized.

Both works use selective inactivation of different subcircuits to explore the temporal dependence of sequential activation of motor activity. Both works also point to the importance of feedback from downstream areas back to high level interneurons and motoneurons in coordinating complex motor sequences. Together these two works pave the way for future experiments exploring motor sequence generation in further detail.

3.7 Manuscript Information

3.7.1 Submitted for publication as

A previous version of this chapter was submitted as [159]:

Quan Wen, Elizabeth Hulme, Sway Chen, Xinyu Liu, Marc Gershow, Andrew M Leifer, Victoria Butler, Christopher Fang-Yen, William R Schafer, George Whitesides, Matthieu Wyart, Dmitri B Chklovskii, and Aravinthan D T Samuel. Bending waves during caenorhabditis elegans locomotion are driven and organized by proprioceptive coupling. *Nature (submitted)*, 2011

3.7.2 The Author's Contribution

The overwhelming majority of the work in this chapter was performed by Quan Wen. Andrew M. Leifer assisted with optogenetic experiments and provided technical support regarding instrumentation and experimental design. Additionally, Leifer made major contributions to the analysis software.

The majority of the manuscript was written by Quan Wen and Aravinthan D.T. Samuel.

Figure 3.5 ***(following page)*****:** Schematic of the motor circuit in C. elegans. (**a**) Worms undulate by alternating contraction and relaxation of dorsal and ventral muscle cells lining the body. Dorsal bending is achieved when dorsal muscle cells contract (filled cells) and ventral muscle cells relax (open cells). Ventral bending is achieved when ventral muscle cells contract and dorsal muscle cells relax. (**b**) General patterns of connectivity in the wiring diagram for forward movement adapted from [75, 142] . Arrows indicate excitatory chemical synapses from the cholinergic motor neurons (VB and DB, in light and dark blue, respectively). Blunt ended lines indicate inhibitory chemical synapses from GABAergic motor neurons (DD and VD, in red and orange, respectively). GABAergic neurons are dispensable for the propagation of the bending wave along the worm body. Dashed lines indicate gap junctions between neighboring muscle cells and neighboring neurons of each cell type. Six to twelve neurons of each cell type are distributed along the worm body.
(**c**) Schematic of the morphology of B-type cholinergic motoneurons that are required for forward movement and A-type cholinergic motor neurons that are required for backward movement. Arrows indicate synapses, circles represent somas and lines indicate the direction and path of neuronal processes. B-type motor neurons extend long processes without synapses in the posterior direction. A-type motor neurons extend long processes without synapses in the anterior direction (adapted from [24]).

Figure 3.5: (continued)

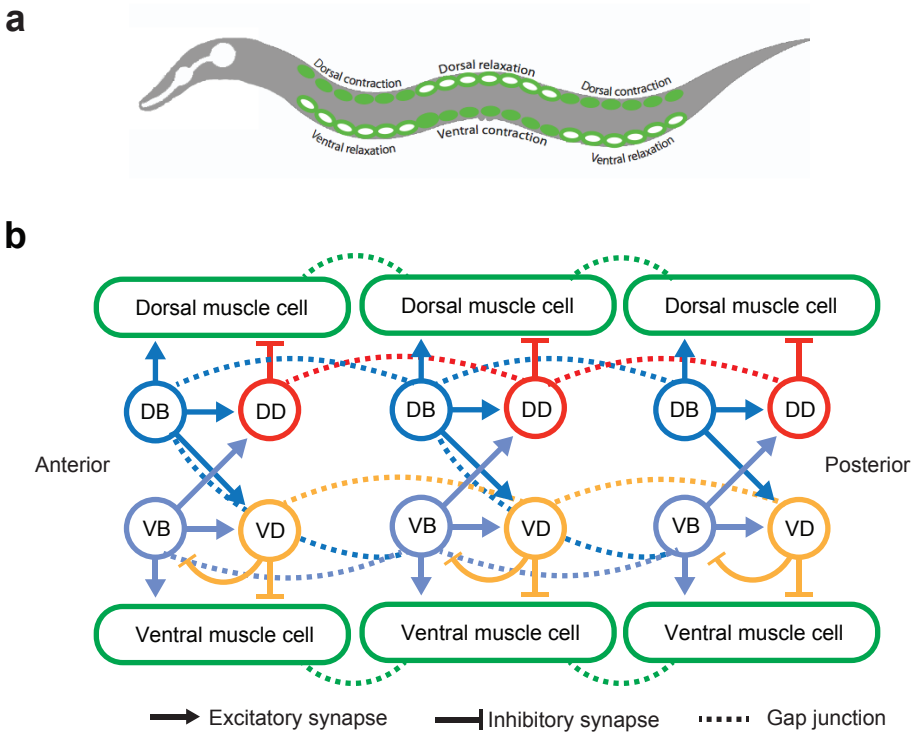


Figure 3.6 ***(following page)*****:** Intracellular calcium dynamics in muscle cells of a transgenic worm (P*myo3::G-CaMP3::RFP*) partially constrained in a straight microfluidic channel.
(**a**) Video images of a worm exhibiting forward locomotory gait while partially immobilized in a straight microfluidic channel. Red fluorescence constitutes the reference channel emanating from RFP in the muscle cells. Green fluorescence constitutes the calcium-sensitive signal emanating from GCaMP3. White lines show the boundary of the channel (also see Video 3.5.2). (**b**) Ratiometric analysis of intracellular calcium dynamics within the ventral and dorsal muscle cells of the worm trapped in the straight channel shown in (**a**). Higher ratios of green to red fluorescence indicate higher intracellular calcium levels. Intracellular calcium levels in both the ventral (upper kymograph) and dorsal muscles (lower kymograph) of the posterior region emerging from the straight channel are lower than in the anterior region. Representative result from one of the five worms studied. (**c**) Kymograph of time-varying body curvature for the worm trapped in the straight channel shown in (**a**).

Figure 3.6: (continued)

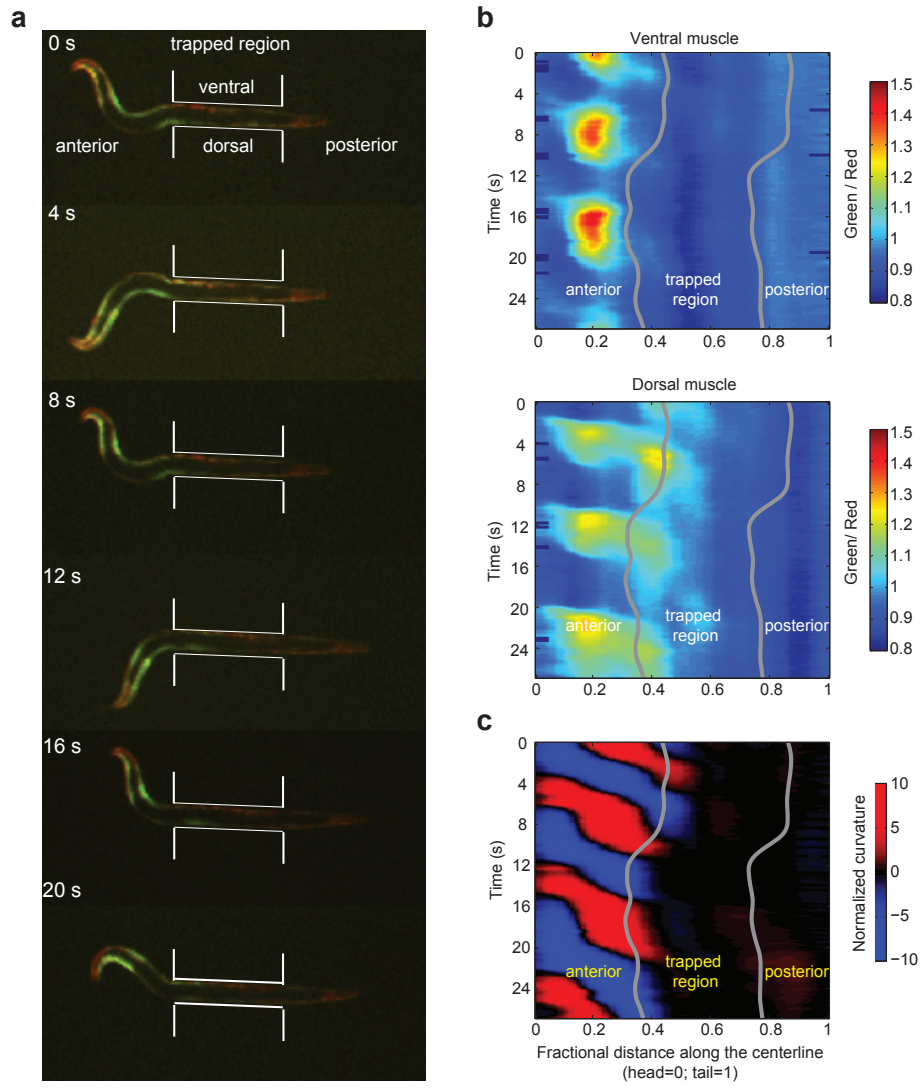
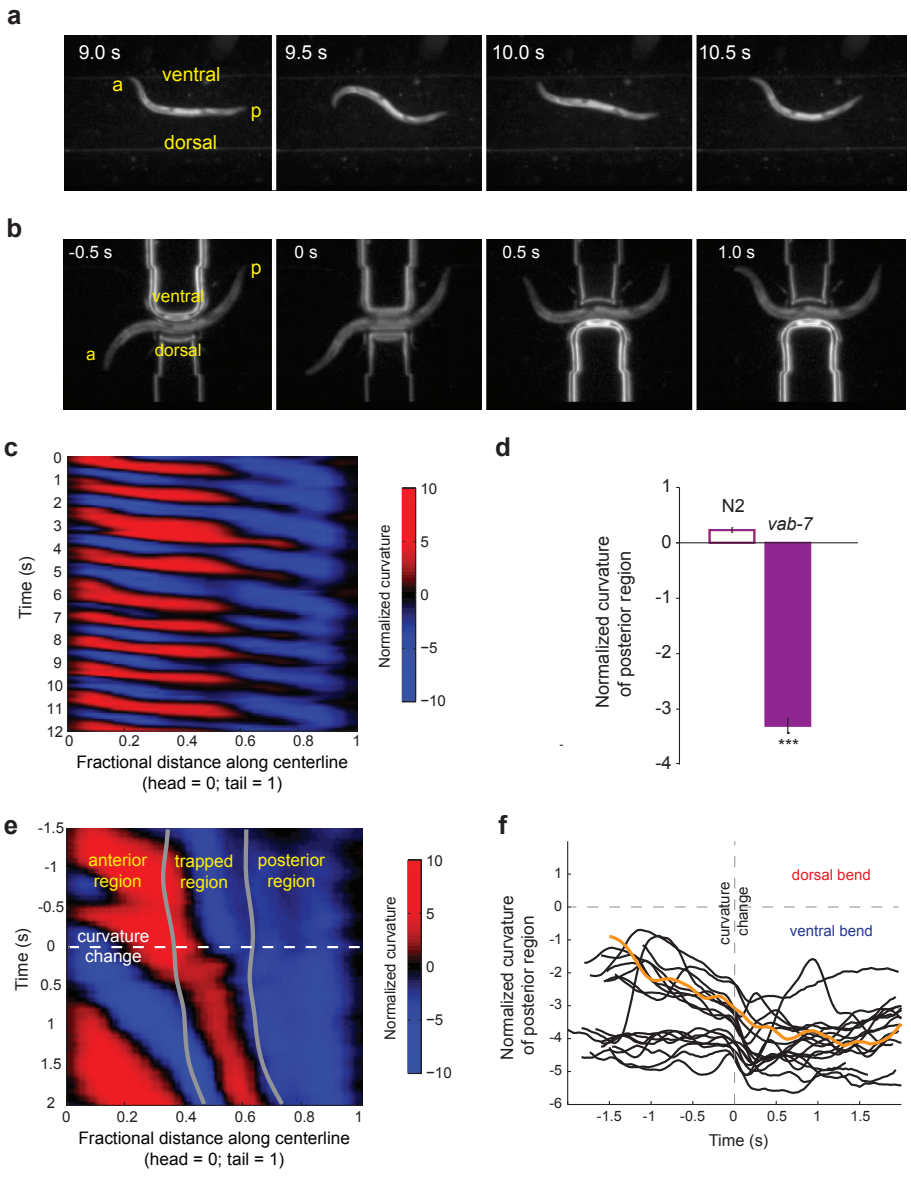


Figure 3.7 *(following page)*: Undulatory wave propagation disrupted in *vab-7* mutants. (**a,b**) Video images of a *vab-7* mutant worm freely swimming in water (**a**) and trapped in a pneumatic microfluidic device (**b**). Ventral and dorsal side of the worm are distinguished by using the eggs as markers. (**c**) Kymograph of a freely swimming *vab-7* mutant worm show bending waves propagate from head to tail. Anterior regions alternate between dorsal curvature (red) and ventral curvature (blue). Posterior regions alternate between null curvature (black) and ventral curvature (blue) (also see Video 3.5.9). (**d**) Curvature of the posterior body region (0.6-0.8 body length), average over an integer number of undulation periods, in a free swimming worm. $n6$ for each data point. $^{***}p = 0.0002$, Mann-Whitney U test. (**e**) Kymograph of *vab-7* mutant worm trapped in the pneumatic microfluidic device as shown in (**b**). Changing the curvature of the trapped middle region towards the dorsal side (red) does not induce dorsal curvature in the posterior region (also see Supplemental Video 10). (**f**) Curvature of the posterior body region, measured as an average from the posterior limit of the channel to the tail, before and after the induced curvature change in the trapped middle region at $t = 0$. Representative trials from six worms. Orange curve corresponds to the experiment shown in (**b**) and (**e**).

Figure 3.7: (continued)



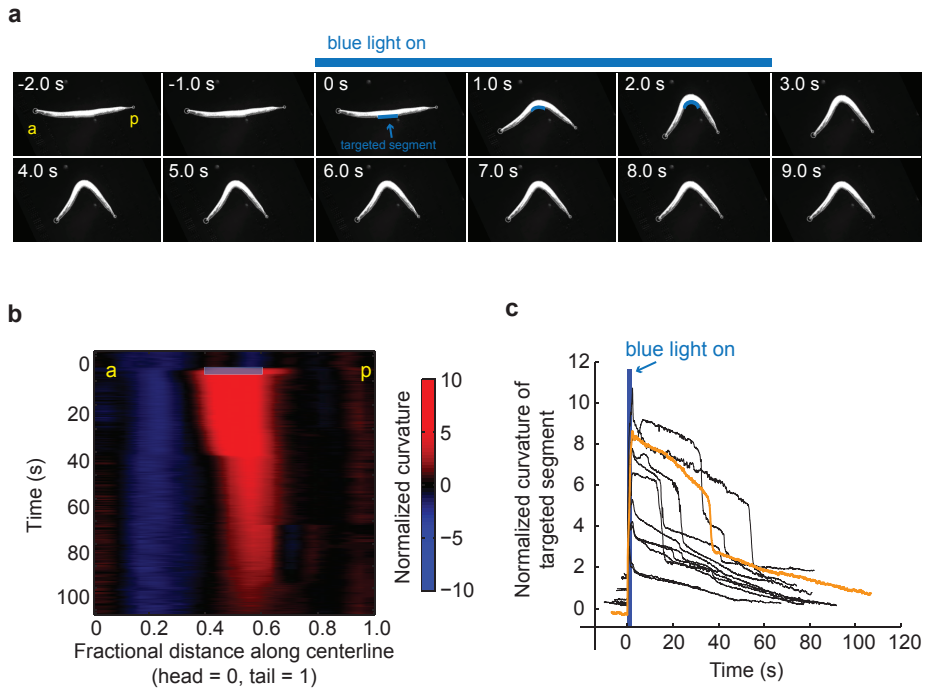


Figure 3.8: Muscle hysteresis induced by optogenetic illumination. (**a**) Video images of an ivermectin-treated transgenic worm (P*myo3::ChR2*). In a 2 s interval, a targeted rectangular region near the middle of the worm (blue rectangle superposed on video images) was illuminated by blue light to induce body bending (also see Video 3.5.11). (**b**) Curvature kymograph of the animal shown in (**a**). White rectangle marks the region and duration of blue light illumination. (**c**) Normalized curvature of the targeted region before and after illumination. Data from three different worms are shown. Orange data points correspond to the experiment shown in (**a**) and (**b**).

4

Sub-micrometer Geometrically Encoded Fluorescent Barcodes Self-Assembled from DNA

4.1 Introduction

In biology and medicine, researchers often use fluorescence microscopy to visualize nanometer- to micrometer-sized entities. In many cases, it is desirable to visualize more than one class of objects simultaneously and unambiguously. Accordingly, there is a need to develop suitable fluorescent tags–or barcodes–for multiplexed imaging applications. Most previously described fluorescent barcodes are constructed using either intensity encoding [51, 63, 96, 99, 105, 112, 166] or geometrical encoding [17, 36, 55, 59, 95, 120, 128, 165]. Intensity encoding relies on the combination of multiple spectrally differentiable fluorophores in a controlled molar ratio. Geometrical encoding, on the other hand, is obtained by separating optical features beyond the microscope's resolution limit (typically ~250

nm for diffraction-limited imaging and ~50 nm for current super-resolution imaging) and arranging them in a specific geometric pattern. Here, we combine both intensity and geometric methods. The multiplexing capability of geometrically encoded barcodes increases exponentially as additional spatially distinguishable fluorophores are incorporated. Therefore, larger barcode libraries may be more easily accessible through geometrical encoding, provided that a rigid structural scaffold capable of defining the spatial arrangement of the fluorescent molecules is available.

To date, despite the remarkable success in synthesizing fluorescent barcodes for in vitro multiplexed detection, very little effort has been made to create robust single-molecule barcodes suitable as *in situ* imaging probes. In addition, most existing fluorescent barcodes range from 2 μm to 100 μm in size, leaving the construction of fluorescent barcodes with the largest dimension less than 1 μm an underexplored research area (with only a few reports [93, 95, 96, 99] and no more than 11 distinct barcodes experimentally demonstrated). Here, we report a group of geometrically and intensity encoded fluorescent barcodes self-assembled from DNA that can be used to tag yeast cells. These barcodes are 400 to 800 nm in length, structurally rigid, biocompatible, reprogrammable in a modular fashion and easy to decode using epi-fluorescence, total internal reflection fluorescence (TIRF) or super-resolution fluorescence microscopy. As evidence of the multiplexing power of the system, 216 distinct barcode species were constructed and resolved using diffraction-limited TIRF microscopy. This is 20 times more than in previous experimentally demonstrated systems.

Structural DNA nanotechnology takes advantage of the well-defined double helical structure of DNA and the highly predictable Watson-Crick base-paring rules to self-assemble designer nano-objects and devices [2, 100, 119, 139, 141**?**]. In recent years, DNA origami has emerged as a prominent method to fabricate two- and three-dimensional structures with sizes of tens to hundreds of nanometers [5, 39, 41, 61, 62, 84, 97, 135]. By folding a long, single-stranded DNA molecule (a scaffold strand, often a M13 viral genomic DNA or its derivatives) with the help of many short synthetic DNA oligonucleotides (staple strands), this approach generates complex, shape-controlled, fully addressable nanostructures. With certain functional groups attached to selected staple strands or their extensions, such nanostructures can be used to organize fluorescent guest

molecules, including small organic molecules [79, 110, 145] as well as metallic [124] and semi-conductive nano-particles [19]. In addition, individual DNA-origami nanostructures can be joined together in a programmable way to make micrometer-sized structures while maintaining their unique nanometer-scale spatial addressability [104, 163]. These properties make DNA origami promising material to build robust fluorescent barcodes, especially as control over the exact ratio of different fluorophores allows intensity encoding while spatial positioning of fluorophores facilitates geometrical encoding and can help minimize undesired inter-fluorophore quenching.

4.2 A Simple DNA Origami Barcode

As a proof-of-concept demonstration, we first designed a family of 27 barcodes based on six-helix bundle DNA nanotubes [40] that are ~800 nm long. Figure 4.1a illustrates the design of such barcodes. Three 84-base pair (~28 nm) zones on the nanotube were selected for fluorescent labeling, with inter-zone distances of 450 nm and 270 nm between the first and the last two zones, respectively. In this design, the fluorescently-labeled zones were separated beyond the diffraction limit of visible light (~250 nm) and the symmetry of the nanotube was broken so that the barcodes could be geometrically encoded with a distinguishable "head" and "tail." For example, labeling the three zones (from left to right in Figure 4.1a) with "Blue" (B, Alexa Fluor 488), "Red" (R, Alexa Fluor 647) and "Green" (G, Cy3) fluorophores (the pseudo-colors were assigned to reflect the excitation wavelength of the fluorophores) resulted in a BRG barcode that should be distinguishable from a GRB barcode. Therefore, a total of $3^3 = 27$ different barcodes can be made from three spectrally distinguishable fluorophores. Within each zone (Figure 4.1b), fluorescently modified oligonucleotides were hybridized to the twelve 21-base long staple extensions protruding out from the main body of the nanotube. The distance between adjacent staple extensions was ~6 nm. To image these prototype barcodes, ten additional staple strands (five per monomer) were designed with 5' biotinylated extensions to enable surface attachment See Supplementary Figure S.1.1 for a strand diagam of the full design.

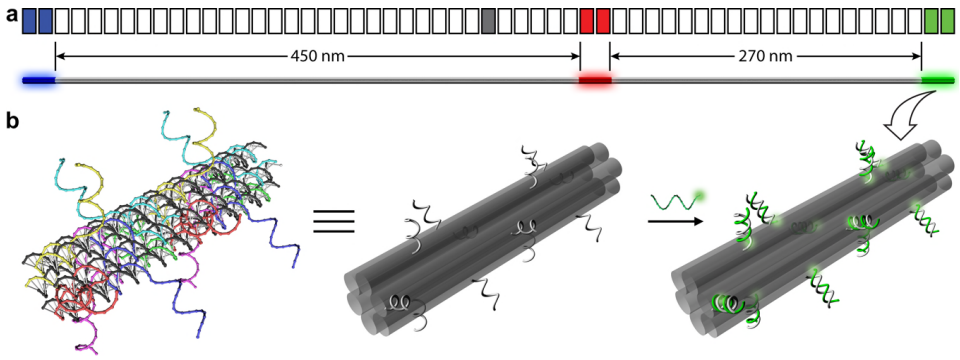


Figure 4.1: Design of the DNA-nanotube-based barcode. (**a**) Two schematic drawings of the Blue–Red-Green (BRG, "–" and "-" denotes larger and smaller inter-zone distance in the barcode, respectively) barcode with a segment diagram on the top and a 3D view at the bottom. The main-body of the barcode is a DNA nanotube formed by dimerizing two origami monomers, each consisting of 28 segments of length 42-bp (13.6 nm). The grey segment in the middle represents the junction where the two monomers are joined together through cross-hybridization between their scaffolds and staples. Three 84-bp zones of the nanotube are fluorescently labeled (shown as blue, red and green segments) to produce the BRG barcode with an inter-zone distance of 450 nm between the first two zones and 270 nm between the last two. Note that each zone is only labeled with one fluorophore species. The resulting barcodes are thus referred as single-labeled- zone barcodes. (**b**) 3D cartoons showing the details of one fluorescently labeled zone. Left: a scaffold-plus-staple model of such an 84-bp zone before labeling. Each of the twelve 63-base-long staples (shown in rainbow colors) contains two parts: the 42-base region at the 5'-end weaves through three double-helices to fold the scaffold (shown in black) into a six-helix bundle nanotube; and the 21-base extension at the 3'-end protrudes out for fluorescent labeling. Middle: an identical but simplified model to emphasize the six-helix bundle structure (each helix shown as a semi-transparent grey cylinder) and the positioning of the twelve 21-base staple extensions (each shown as a light-grey curl). Right: Cartoon representation of a "green" 84-bp zone. The labeling is achieved by hybridizing the Cy3 (shown the glowing green spheres at the 3'-ends) modified strands to the staple extensions.

4.2.1 Construction of Simple Barcode

We assembled the DNA-origami nanotubes following the protocol in [40] with slight experimental modifications (see Materials and Methods in SI). Briefly, two structurally identical but chemically distinct nanotube monomers (~400 nm each) were assembled in two separate test tubes by slowly cooling a mixture of 7.3-kb scaffold strands and a set of ~200 staple strands from 80 °C to 24 °C over 15 hours. After folding, excessive staple strands were removed from the folded nanotubes through polyethylene glycol fractionation. The nanotube monomers were then incubated with the desired fluorescently-modified oligonucleotides and the labeled monomers were mixed at an equimolar ratio to form the final barcode structure. The barcodes were subsequently purified via agarose-gel electrophoresis and immobilized on streptavidin-coated coverslips before imaging with epi-fluorescence (Supplementary Figure S.1.2) or TIRF microscopy (Figure 4.2 and 4.3) in the presence of oxygen scavenging reagents [97].

4.2.2 Inspecting the simple barcodes

In order to validate our system, we randomly chose five distinct barcodes from the 27 members in the barcode family for quality control experiments. The barcodes were assembled and purified separately and imaged under the same experimental conditions. Two distinct features of the barcode were clearly visible from the TIRF images (Figure 4.2a, top panel and Supplementary Figure S.1.3). First, each fluorescently-labeled zone on a barcode was resolved as a single-color spot, and each complete barcode consisted of three such spots. Second, two of the neighboring spots were separated by a small gap while the other two neighbors sat closely together. One can visually recognize and decode those geometrically-encoded barcodes based on the color identity of the spots and their relative spatial positions, even without the aid of specialized decoding software. Using custom-written software that localizes the center of each spot on the BRG barcodes (Supplementary Figure S.1.4), we measured the average center-to-center distance between the neighboring spots to be 433 ± 53 nm (mean$\pm$s.d., N=70; larger distance) and 264 ± 52 nm (mean $\pm$ s.d., N=70; smaller distance), confirming the correct for-

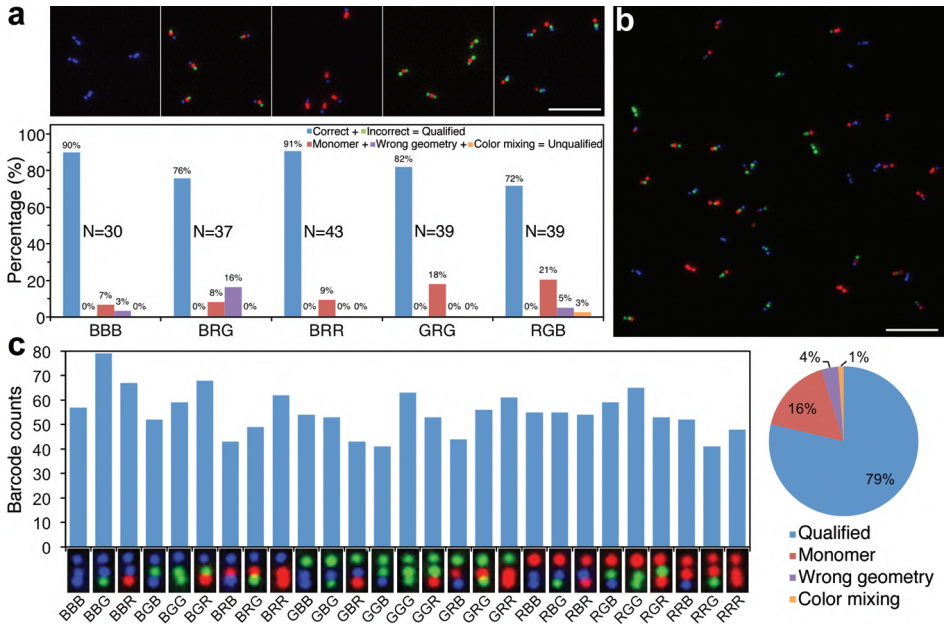


Figure 4.2: Single-labeled-zone fluorescent barcodes. (**a**) Superimposed TIRF microscopy images of five barcode species (top) and the statistics from manual counting (bottom). From left to right are the BBB, BRG, BRR, GRG and RGB barcodes with a representative image on top of the corresponding bar-graph. Each bar-graph is generated based on the manual sorting and counting of the objects found in a 50×50 μm² image (~40 barcodes, the exact sample size N is noted beside the corresponding bar-graph). (**b**) A representative image of the equimolar mixture of 27 barcode species. (**c**) Statistics obtained by analyzing twenty-seven 50×50 μm² images of the 27 barcode mixture (~1,500 barcodes in total). Left: barcode counts of the 27 species (average count of 55 with a standard deviation of 9). A representative TIRF image (1.4×0.7 μm²) of each barcode type is placed underneath the corresponding bar. Right: sorting result of the observed objects shown as a pie-chart. Color scheme used for the bar-graphs and the pie- chart (unrelated to the pseudo-colors of the fluorophores): blue, correct barcodes (qualified barcode with expected identity); green, incorrect barcodes (qualified barcode with unexpected identity); red, monomer nanotubes (one spot or two connecting spots); purple, barcodes with wrong geometry (i.e., bending angle <120°, see methods in SI); and orange, barcodes containing at least one spot with two colors. Note that in the 27- barcode pool, correct vs. incorrect barcodes were not distinguishable because all barcode types are expected. As a result the bars and pie representing the qualified barcodes in (**c**) are shown in blue. Scale bars: 5 μm.

mation of the barcodes. These experimentally measured distances were slightly smaller than the designed values (478 nm and 298 nm). We attribute this discrepancy to random thermal bending of the nanotubes (persistence length of ~1–2 μm), which has been observed previously by others [61, 135] and confirmed by us (Supplementary Figure S.1.5) using transmission electron microscopy (TEM). It is important to note that unlike some other geometrically-encoded barcoding systems (e.g., NanoString nCounter [55]), there was no molecular combing involved in sample preparation. The separation between the fluorescent spots was exclusively created by the inherently rigid structure of the six-helix bundle nanotube. It is also notable that the spot intensities were not perfectly uniform across the whole image, which can be explained by factors such as the uneven illumination of the sample stage and differences in labeling efficiency. Nevertheless, the TIRF images proved that the barcodes were successfully assembled and can be resolved unambiguously.

4.2.3 Quantification of the yield of the simple barcodes

We subsequently manually investigated TIRF images with an area of 50×50 μm² for the five selected barcodes (Figure 4.2a, bottom panel). The objects found within the images were first sorted into qualified (i.e., three single-color spots arranged in a nearly linear and asymmetric fashion as designed) and unqualified (i.e., all other objects) barcodes. The qualified barcodes were further categorized into correct and incorrect (false-positive) barcodes based on the fluorescent signatures of the composing spots to reflect whether the barcode was the expected type. The unqualified barcodes were further sorted into (1) monomer nanotubes (single spot or two "kissing" spots), (2) barcodes with "wrong" geometry (i.e., extreme bending), and (3) barcodes containing at least one spot with multiple colors. Our statistics revealed that more than 70% of the visible objects were qualified barcodes, which we further determined to be exclusively the expected type (i.e., zero false-positive out of 188 qualified barcodes observed). The unqualified barcodes likely arose from folding defects, sample damage during handling and overlapping nanotubes on the surface. These errors can be reduced by optimizing the sample preparation and imaging protocol.

To be useful for multiplexed imaging system, a number of different barcode

species must coexist in one pool. Thus, it is important to examine the robustness of our system by mixing different types of barcodes together. In an initial test, we synthesized BRG and RGB barcodes separately, mixed them together at equal molar ratio and co-purified them via gel electrophoresis. The TIRF analysis of the purified mixture (Figure S.1.6) confirmed the 1:1 stoichiometry of the two barcodes and the overall assembly success rate (qualified barcode/all objects) of ~80%, suggesting that both barcodes maintained their integrity in the mixing and co-purification process. In addition, over 98% of the qualified barcodes fell into one of the two expected types (BRG and RGB). The 2% false-positive rate was due to an unexpected barcode, namely BGB (Supplementary Figure S.1.6), which could be attributed to a rare occasion in which the front monomer of the BRG barcode lays in proximity to the rear monomer of the RGB barcode. We believe this can be eliminated in the future if a more stringent purification condition is applied to minimize the amount of leftover monomers.

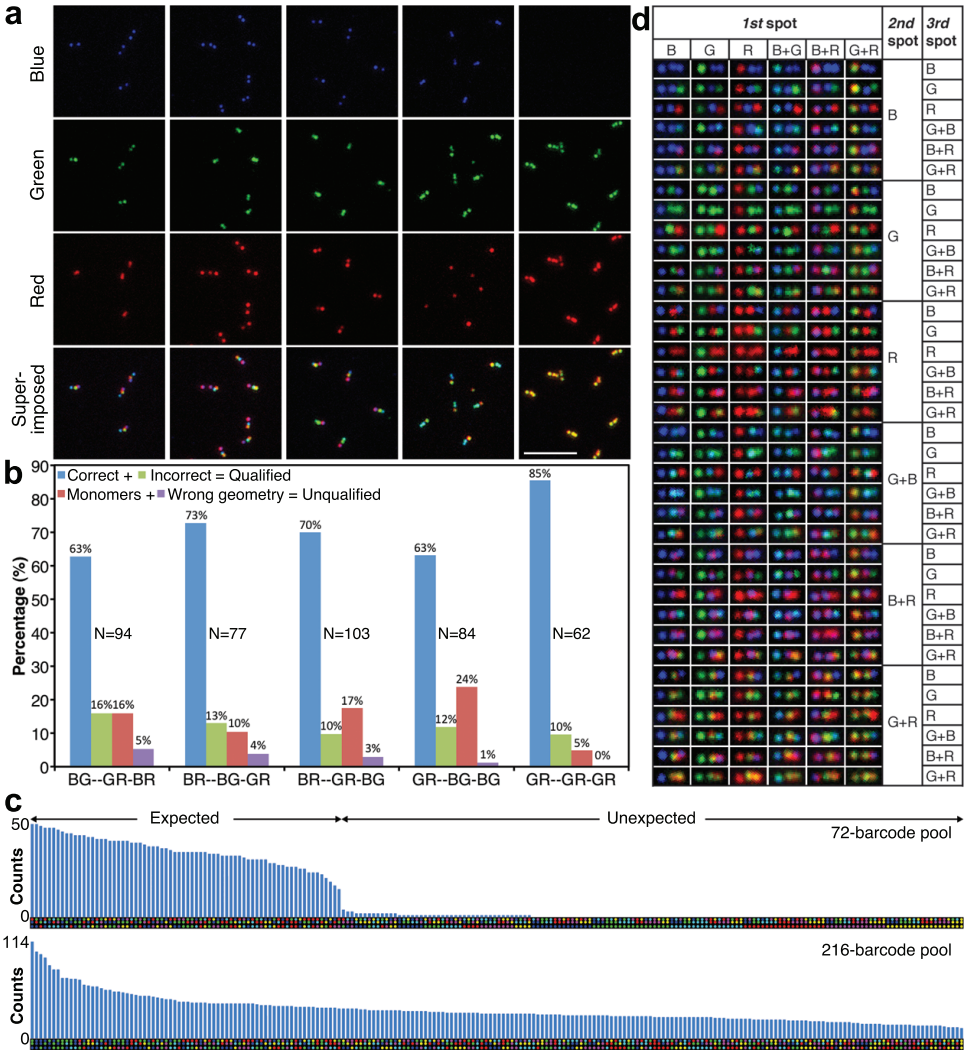
We next inspected barcodes when all 27 members of the barcode family were mixed at equimolar amount. The TIRF images (Figure 4.2b and Figure S.1.7) showed that all types of barcodes were resolved. Statistical analyses of 27 50×50 μm^2 images (~1,500 barcodes in total) revealed an average count of 55 per barcode type with a standard deviation of 9 (Figure 4.2c), fitting well with the expected stoichiometry given pipetting and sampling errors. The distribution of observed objects over the four categories (note that here correct vs. incorrect barcodes were not distinguishable as all 27 types were included) was consistent with the values measured from the single-type barcode samples. The above observations suggest that the sub-micrometer-long DNA nanotube represents a reliable platform to construct geometrically-encoded barcodes with built-in structural rigidity.

4.3 Dual-labeled DNA origami Barcodes

Our modular design enabled straightforward reengineering of the barcode to enhance its multiplexing capability through a variety of ways. For example, we changed the sequence of six staple extensions per zone. Instead of using 12 identical fluorescent oligonucleotides for labeling, we had the ability to use a combination of up to two fluorophores to create more unique fluorescence signatures

Figure 4.3 ***(following page)*****:** Dual-labeled-zone fluorescent barcodes. (**a**) Typical TIRF microscopy images of five selected barcode species, shown both in separate channels and after superimposing. Scale bar: 5 μm. (**b**) Statistics obtained by analyzing two 50×50 μm² images of each barcode species (~85 barcodes, the exact sample size N is noted beside the corresponding bar-graph). The barcode types are noted under the x-axis of the diagram. Color scheme (unrelated to the pseudo-colors of the fluorophores): blue, correct barcodes (correct geometry and color identity); green, incorrect barcodes (correct geometry but incorrect color identity); red, monomer nanotubes (one spot or two connecting spots); and purple, barcodes with wrong geometry (i.e., bending angle <120°, see methods in SI). (**c**) Computer-aided barcode counting results of the 72-barcode pool (N=2,617) and the 216-barcode pool (N=7,243) plotted as bar-graphs with descending barcode counts from left to right (see Supplementary Tables S.1.6 and S.1.7 for the numeric counting results). A computer-generated reference barcode image is placed underneath the corresponding bar. (**d**) A table containing one representative TIRF image (1.4×0.7 μm²) for each of the 216 dual-labeled-zone barcode species.

Figure 4.3: (continued)



(pseudo-colors) for each zone. Six pseudo-colors (B, R, G, BG, BR, and GR) were generated by this "dual-labeling" strategy using three spectrally differentiable fluorophores. Consequently, the total number of distinct barcodes was raised from $3^3 = 27$ to $6^3 = 216$, an order of magnitude increase in the multiplexing capability.

4.3.1 Inspecting the dual-labeled barcodes

Similar to the single-labeled-zone barcode family, five members from the dual-labeled-zone barcode family were chosen for quality control purposes. The barcodes can be visually decoded either solely from the superimposed image or by examining all different channels simultaneously. For example, as shown in the first column of Figure 4.3a, the barcode "BG–GR-BR" ("–" and "-" denotes larger and smaller inter-zone distance in the barcode, respectively) exhibits two spots each in the blue, green and red channels but with descending gaps between them, matching its design. In the superimposed image, the barcodes are seen as Cyan–Yellow-Pink, an expected consequence of color mixing caused by the dual-labeling strategy. In similar fashion, we further verified the correct formation of the other four selected barcodes (Figure 4.3a and Supplementary Figure S.1.8). Although the final pseudo-color from the dual-labeled zones was not always uniform (e.g., some yellow spots were green-tinted while the others were red-tinted) due to inconsistent labeling efficiency and minor sample displacement during imaging, the fluorescence signature of any given spot could be identified by checking the raw images acquired from the three imaging channels. We manually analyzed two 50×50 μm^2 images of each dual-labeled-zone barcode and plotted the statistical data in Figure 4.3b. Here, objects were sorted into qualified barcodes and unqualified barcodes based on their geometry. The qualified ones were further categorized as either correct or incorrect. We found that 75 to 95% of the objects were qualified barcodes, among which 80–90% were the correct type (percentage varies depending on the exact type of barcode). Compared to the single-labeled-zone barcode family, the percentage of qualified barcodes remained the same, while the false positive rate increased significantly from zero (Figure 1a) to 10–20.

We further tested the dual-labeled-zone barcoding system by imaging a mixture containing 72 barcode species that were individually assembled and co-purified

(Supplementary Figure S.1.9). Custom MATLAB scripts were used to assist the decoding process (See Methods, Section 4.9.9) in two steps. In step one, a three-channel (red, green, blue) TIRF image containing barcodes was pre-processed to remove background and thresholded so that only pixels containing qualified barcodes remained. The resulting three-channel binary image was merged to generate a single-channel binary image. Next, the software identified the location and orientation of geometrically-legitimate barcodes based on their shape in the binary image. In step two, for each barcode located in step one, the corresponding region of the three-channel image was compared against a library of all possible reference barcodes. The observed barcode was assigned the identity of the reference barcode with the highest correlation. The fully-automated decoding process (unsupervised mode) ended after the above two steps. In an optional supervised mode, the software presented the user with the observed barcode and its most likely identity for approval. Comparison between supervised and unsupervised decoding results confirmed >80% agreement between the computer and the user (Supplementary Figure S.1.11). The computer-aided (supervised mode) analysis of thirty six 64× 64 μm² three-channel images registered ~2,600 qualified barcodes that belonged to 116 different species (Table S.1.6 and Figure 4.3c, top panel). The expected 72 species constituted ~98% of the total barcode population with an average barcode count of 36 per species and a standard deviation of 8. In contrast, the unexpected species averaged only ~1.4 barcodes per species (maximum 4 counts). Finally, we analyzed a mixture containing all the 216 members of the dual- labeled-zone barcode (Supplementary Figure S.1.12). Sixty 64× 64 μm² images of this mixture were processed by the decoding software in the unsupervised mode. The fully-automated analysis registered a mean barcode count of ~34 per species (~7,200 barcodes total) with a standard deviation of 17 (Supplementary Table S.1.7 and Figure 4.3c, bottom panel). Our study demonstrates that 216 barcode species were successfully constructed and resolved (Figure 4.3d).

4.4 DNA origami barcode system requiring super-resolution microscopy

Increasing the number of spatially differentiable fluorescently-labeled zones on the origami nanotube is another way to enhance the multiplexing capability of the

barcoding system. A straightforward solution that keeps the inter-zone distances larger than the diffraction limit would require longer (e.g. a few micrometers long) nanotubes, which would likely result in lower assembly yield and decreased mechanical rigidity. Alternatively, one can add more fluorescently-labeled zones on the 800 nm tube and image the barcode using super-resolution fluorescence microscopy. As a feasibility demonstration of the latter approach, we applied DNA-PAINT [79], a recently developed super-resolution fluorescence technique, to image the barcodes.

4.4.1 Introduction to super-resolution microscopy techniques

In recent years, several techniques have been developed allowing imaging beyond the diffraction limit using far-field fluorescence microscopy [69, 70, 74, 155, 158]. In most super-resolution implementations, fluorophores are switched between fluorescence ON and OFF states, so that individual molecules can be localized consecutively. In methods relying on targeted readout schemes such as in Stimulated Emission Depletion Microscopy (STED) [71] or other Reversible Saturable Optical Fluorescence Transitions (RESOLFT) techniques [69], fluorescence emission is actively confined to an area below the diffraction limit. The switching of fluorescent molecules can also be carried out stochastically such as in Stochastic Optical Reconstruction Microscopy [68, 136] (STORM, dSTORM), Photoactivated Localization Microscopy [11] (PALM) and Blink Microscopy (BM) [144], where most fluorescent molecules are "prepared" in a dark state and only stochastically switched on to emit fluorescence. In Point Accumulation for Imaging in Nanoscale Topography (PAINT) [140], fluorescence switching is obtained by targeting a surface with fluorescent molecules. In all stochastic approaches, fluorescence from single molecules is localized in a diffraction-limited area to yield super-resolved images [169, 170].

DNA-PAINT uses transient binding of fluorescently-labeled oligonucleotides (imager strands) to complementary "docking" strands on DNA nanostructures to obtain switching between a fluorescence ON and OFF state, which is necessary for localization-based super-resolution microscopy (See Figure 4.4a). By adjusting the length of the imager/docking strand duplex and the concentration of imager strands in solution, fluorescence ON- and OFF-times can be tuned independently.

4.4.2 DNA-PAINT-based DNA origami barcodes

For this study, we extended the DNA-PAINT technique to three-color imaging using orthogonal imager strand sequences coupled to three spectrally-distinct dyes (Atto488 for blue, Cy3b for green and Atto655 for red excitation). To demonstrate the feasibility of the three-color super-resolution barcode system, we designed a DNA nanotube monomer with four binding zones in a symmetric arrangement. The neighboring zones were separated by ~114 nm (i.e. well below the diffraction limit). Each binding zone consists of 18 staple strands, which can be extended to display three groups of orthogonal sequences (six per group) for the red, green or blue imager strands to bind. As a proof-of-principle experiment, we designed five different barcodes (Figure 4.4a and top panel of b). The bottom panel of Figure 4.4b shows the super-resolution reconstruction of the five barcodes for each channel separately as well as an overlay of all channels. Figure 4.4c shows a larger area containing all five barcodes. The unique pattern of the barcodes in all three channels can be resolved. Some of the barcodes can move during the sequential imaging (100 s acquisition time per channel) of all three color channels (see overview image in Supplementary Figure S.1.14). Image conditions could still be improved by alternating excitation and faster image acquisition to prevent this effect. The transient, repetitive binding of imager strands to docking sequences on the nanotube not only creates the necessary "blinking" behavior for localization, but also makes the imaging protocol more robust, as DNA-PAINT is not prone to photobleaching or incorrectly labeled strands (Supplementary Figure S.1.13). With the microscope setup we used, DNA-PAINT provides a resolution of ~27 nm (FWHM of a Gaussian fit to the reconstructed PSF) in the red, ~22 nm in the green and ~23 nm in the blue channel. The obtainable resolution and imaging specificity suggests that 6 positions on one nanotube monomer could be robustly resolved while keeping the geometrical asymmetry of the barcode, which would lead to $7^6 = 117,649$ different possible barcodes. Furthermore the modularity of the nanotube design enables the customized reengineering of barcodes with inter-zone distances tailored to the resolving power of the used microscope, thus making it applicable for a wide range of microscope setups.

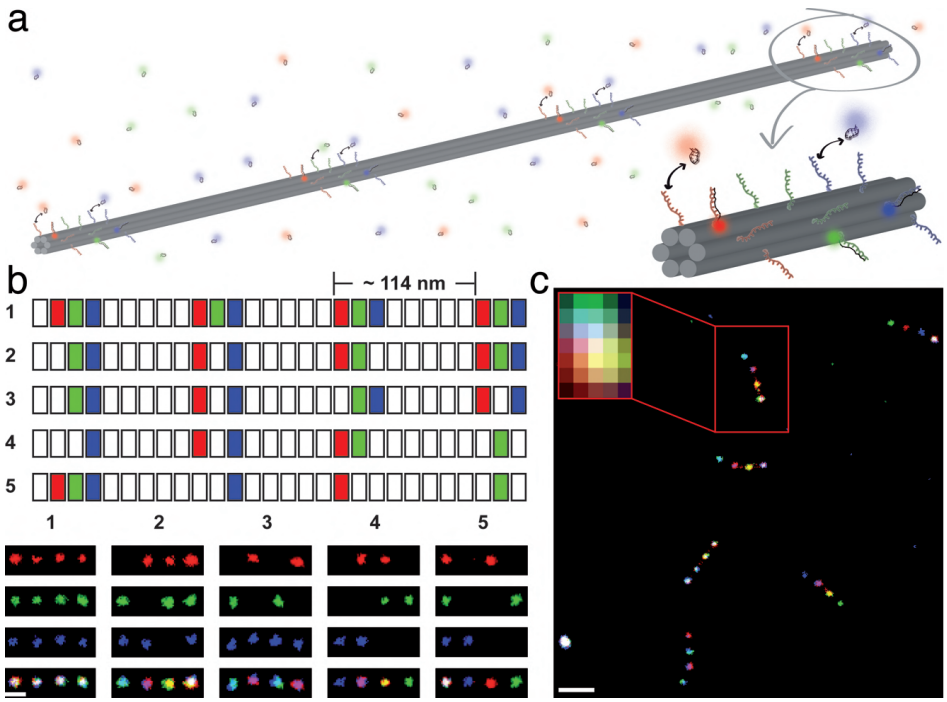


Figure 4.4: Super-resolution fluorescent barcodes. (**a**) Scheme of DNA-PAINT used for super-resolution barcode imaging. The 400 nm nanotube consists of 4 binding zones spaced by ~114 nm. Each zone can be decorated with the desired combination of "docking" sequences for red, green or blue imager strands. The orthogonal imager strands bind transiently to their respective "docking" sites on the nanotube, creating the necessary "blinking" for super-resolution reconstruction (**b**) Top: segment diagram (similar to the one used in Figure 4.1) of the DNA nanotube monomers used for creating five barcodes for super-resolution imaging. Bottom: super-resolution images of the five barcodes shown in each channel separately and as an overlay of all channels. Scale bar: 100 nm. (**c**) Super-resolution image showing all five barcodes in one mixture. The inset shows the diffraction-limited image of one of the barcodes. Scale bar: 250 nm.

4.5 More Complex Geometries

DNA nanostructures with non-linear geometry could be assembled to generate more sophisticated barcodes. Figure 5 shows how three ~400 nm DNA tubes were linked to the outer edge of a ~60 nm DNA ring through hybridization between staple extensions (Figure 4.5a, inset). Fluorescently labeling the ring and the far end of the nanotubes generated a three-point-star-like structure clearly resolvable under fluorescence microscopy. TIRF microscopy and TEM studies (Figure 4.5b, Supplementary Figures S.1.15 and S.1.16z) revealed that about 50% of successfully folded barcodes featured three nanotubes surrounding the ring with a roughly 120° angle between each other as designed, while many other barcodes had significantly biased angles between neighboring nanotubes due to the semi-flexible double-stranded DNA linker between the ring and the nanotubes. It is conceivable that by using a similar design to connect three identical "satellite" linear barcodes to a central hub (here the three satellite barcodes may share the hub as a common fluorescently labeled zone), one can construct barcodes with triplicated encoding redundancy with reliability. In addition, more rigid linkers between the ring and the protrusions (e.g. multi-helix DNA with strand crossovers) can be employed to enforce better-defined barcode geometry.

4.5.1 DNA-PAINT-based DNA origami barcodes

4.6 DNA Barcodes attached to cells

Modifying barcodes with functional ligands such as antibodies and aptamers would allow them to tag specific biological samples and serve as multiplexed *in situ* imaging probes. In a proof-of-principle experiment, the GRG barcode was used to tag wild-type *Candida albicans* yeast. The yeast cells were first mixed with a biotinylated polyclonal antibody specific to *C. albicans*, then coated with a layer of streptavidin and incubated with biotinylated GRG barcodes (Figure 4.6a). TIRF microscopy revealed the barcodes attached to the bottom surface of the yeast cells (Figure 4.6b, top panel and Supplementary Figure S.1.17). While some of the nanotubes landed awkwardly on the uneven cell walls of the yeast cells, a number of GRG barcodes can be clearly observed. In contrast, no barcode tagging was

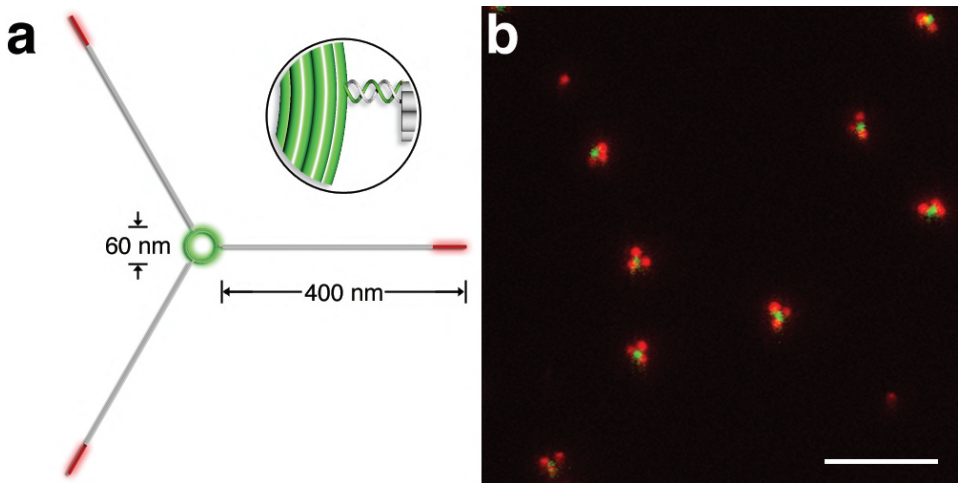


Figure 4.5: Fluorescent barcode with non-linear geometry. (**a**) Schematic. Three identical ~400-nm long DNA nanotubes are linked to the outer edge of a DNA ring with diameter of ~60 nm through the hybridization between staple extensions. The ring and the end of the tube are labeled by Cy3 (green) and Cy5 (red), respectively. (**b**) A representative TIRF microscope image of the barcode shown in (**a**). Scale bar: 5 μm.

observed when non-biotinylated antibodies or barcodes were used to treat the yeasts (Figure 4.6b, bottom panel and Supplementary Figure S.1.17), suggesting that little to no non-specific interaction existed between the barcode and the cell surface.

4.7 Discussion

In summary, we have constructed a new kind of geometrically-encoded fluorescent barcode using DNA origami-based self-assembly. We took advantage of the mechanical rigidity of the multi-helix-bundle architecture, the addressable surface and spatial organizing power of the DNA-origami nanotube, as well as the programmable linkage between multiple DNA nanostructures to build a modular, robust yet easy-to-handle barcoding system. It differs from previously reported optically decodable barcoding systems in the following three aspects: First, and most notably, the DNA origami-based barcodes are 400 to 800 nm in length, making them potentially useful to serve as *in situ* single molecule imaging probes to tag sub-cellular entities such as surface protein markers (as demonstrated in the yeast tagging experiment), chromosomes, mitochondria and microtubules, which generally range from 500 nm to 20 μm in size. Making use of the super-resolution barcode system, one could envision *ex* and *in situ* imaging of a wide variety of biomolecules such as mRNA or proteins simultaneously by using tagging techniques based on hybridization or SNAP tags [54, 78, 85, 87]. The system also fulfills a technological challenge to build robust finite-size optical barcodes with its smallest feature approaching or even smaller than the diffraction limit of visible light.

Second, the barcodes presented here are massively self-assembled in parallel (typically 1011 molecules per barcode species in one 50 μL folding reaction) exclusively through DNA hybridization. Unlike previous constructions, the synthesizing and purification processes require no enzymatic reactions [55], genomic engineering [105], photolithography [128], electrochemical etching [35] or microfluidic devices [51, 128]. This makes the barcodes easy to reproduce in a typical biochemistry lab. The barcodes could be functionalized to display sequence-specific probes simply through DNA-origami staple extensions. Third, the robustness and

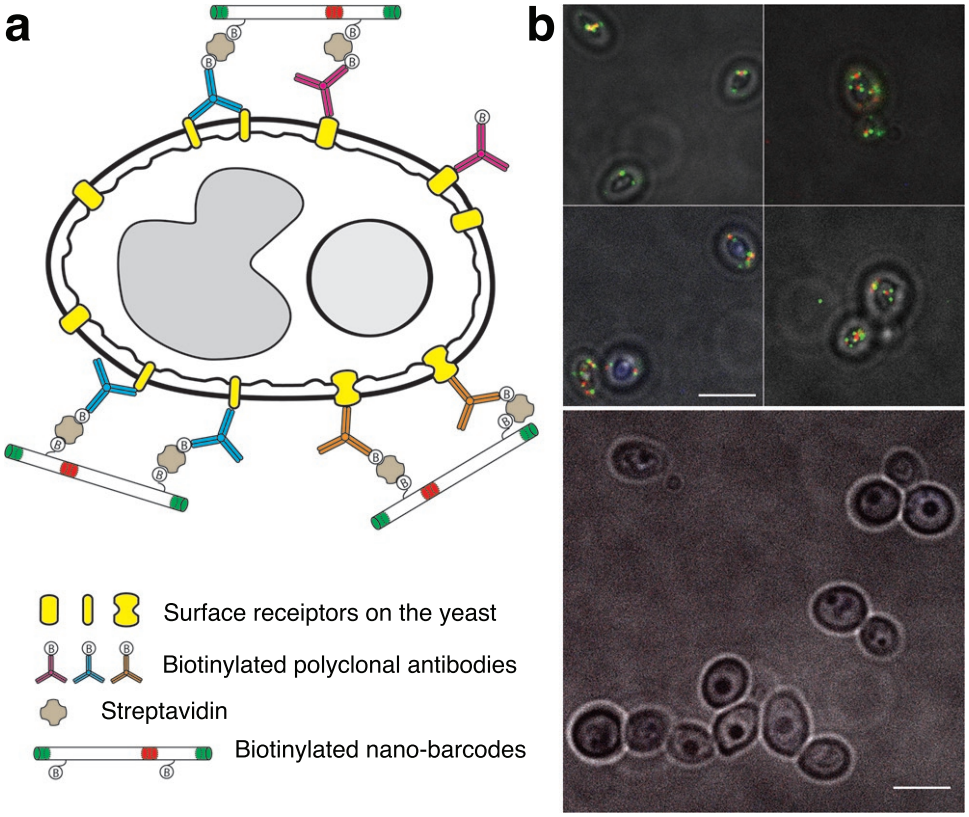


Figure 4.6: Tagging yeast cells with the GRG barcodes as *in situ* imaging probes. (**a**) Cartoon illustrating the tagging mechanism. The biotinylated barcodes are anchored on the yeast cell through streptavidin molecules bound to biotinylated polyclonal antibodies coated on the yeast surface. Only two of the ten biotinylated staples on the barcode are shown here for clarity. (**b**) Overlaid microscope images (acquired in bright field and TIRF) of the yeast cells treated with the barcodes. Top: yeast cells treated as illustrated in (**a**). Bottom: negative control: yeast cells treated with non-biotinylated barcodes. Scale bars: 5 μm.

programmability of the DNA origami-based platform allows us to achieve high multiplexing capability.

Here, we successfully constructed 216 barcode species that were decoded in one pool, while the previously reported sub-micrometer-sized fluorescent barcodes contained no more than 11 distinct barcodes [51, 93, 95, 96]. Even compared to the prior art of producing multi-color fluorescent barcodes that are less than 10 μm in size, our system is among the few examples [51, 55] that demonstrated more than 100 codes in practice. To the best of our knowledge, the only system that demonstrated larger numbers of barcodes coexisting in a homogenous solution is the NanoString nCounter [55]. However, the reporter probes of the NanoString (~7 kb double-stranded RNA-DNA-hybrid molecules) require combing to make the codes resolvable, making them fundamentally unsuitable as *in situ* imaging probes. In contrast, the DNA origami-based barcodes maintain their structural integrity when applied to cell surfaces, suggesting their promising application as *in situ* imaging makers. Furthermore, the combination of the DNA-PAINT imaging technique and the DNA-origami-based nanotubes allowed us to create the first nanoscopic "super-resolution" barcodes while maintaining multiplexing power.

To give an idea of the number of achievable different codes consider the incorporation of 12 fluorescently-labeled zones into the 800-nm long dimer nanotube and using all possible equimolar combinations of the RGB fluorophore. An enormous number ($7^{12} \approx 1.4 \times 10^17$ of barcode species can be generated. Using DNA-PAINT as a readout technique also makes the system less prone to fluorescent labeling errors, as the "imager strands" are replenished from solution during the acquisition process. Finally, the fact that the barcodes can be resolved using epi-fluorescence microscopy makes them more adaptive and implies their possible application in deep tissue labeling in the future.

4.8 Future Directions

We note that barcode design and experimental procedures can be refined to minimize the false-positive rate of the dual-labeled-zone barcodes. For example, fluorophores with higher quantum efficiency and less susceptibility to photobleaching (e.g. Atto dyes or quantum dots) could be used to improve imaging quality.

Higher quality staples (e.g., enzymatically produced and HPLC purified) could be used to fold the barcode to enhance fluorescent labeling efficiency by minimizing the number of missing or truncated staple extensions in the binding zones. Barcodes could be purified under more gentle conditions (e.g. recover DNA from agarose gel through electro-elution [8] instead of crush-and-soak) to reduce the possibility of breaking a dimer nanotube into two monomers. The purified barcodes could be chemically cross-linked [129] to improve their thermal and mechanical stability. Redundant coding could be built into the barcode to generate more reliable readouts. In practice, it is likely that a small number of distinct codes could be sufficient to label the targets of interest. For those applications, one could selectively use the most robust subset of the barcode library. A larger library of barcodes may be achieved through more sophisticated geometrical and intensity encoding, which is made possible by our growing ability to build more complicated DNA nanostructures and to optically resolve features below the diffraction limit. As computational power and image-processing technology advance, future barcode analysis systems should run more accurately. In the future, we expect our barcode technology to interface with the rapidly growing protein/antibody library technology to yield a powerful imaging toolbox for single-molecule studies of biological events and biomedical diagnostics.

4.9 Materials and Methods

4.9.1 Materials

Non-modified DNA Oligonucleotides were purchased from Bioneer Inc. (Alameda, CA). Fluorescently-modified DNA oligonucleotides were purchased from Integrated DNA Technology (Coralville, IA) and Biosynthesis (Lewisville, TX). Polyclonal antibodies to *C. albicans* were purchased from Thermo Scientific (Rockford, IL). Streptavidin was purchased from Invitrogen (Carlsbad, CA). Glass slides and coverslips were purchased from VWR (Radnor, PA). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO). M13mp18 scaffold strand for the DNA origami drift markers was obtained from New England Biolabs (Ipswich, MA).

4.9.2 Design of the nano-barcodes

The main-body of the linear nano-barcode is a DNA six-helix-bundle (6hb) nanotube dimer designed using the caDNAno software (`http://cadnano.org`). The structure of the barcode is illustrated as a strand diagram in Figure S.1.1. Staples in the fluorescent labeling zones were extended at the 3'-end with 21-base long single-stranded overhangs (handles) for fluorescent labeling (see table S.1.1 for staple extension sequences). In addition, ten staples (five per monomer, shown as magenta strands in the diagram) were extended at the 5'-end with biotin. The three-point star-like barcode is self-assembled from three identical DNA 6hb nanotube monomers and a DNA 6hb ring. Here, 12 handles protrude out from a helix on the outer edge of the ring. Three of them serve as sticky-ends to link with the 6hb nanotubes while the other nine are used for fluorescent labeling. The 6hb nanotube carries a handle at the front end to serve as the complementary sticky-end and 24 handles at the rear end for fluorescent labeling. The design is illustrated as a strand diagram in Figure S.1.1, and the staple extension sequences used in this design are listed in table S.1.2.

4.9.3 Self-assembly of the nano-barcodes

To fold the ~800 nm DNA nanotube based barcode, the monomer nanotubes were each assembled and fluorescently labeled in a separate test tube and then mixed together to form the dimer. The assembly of the monomer was accomplished in a one-pot reaction by mixing 100 nM scaffold strands (7,308-base long, termed p7308, see table S.1.4 for sequence) derived from M13 bacteriophage with a pool of oligonucleotide staple strands (600 nM of each; reverse-phase cartridge purified.) in folding buffer containing 5 mM Tris, 1 mM EDTA, 20 mM $MgCl_2$, 50 mM NaCl (pH 8) and subjecting the mixture to a thermal-annealing ramp that cooled from 80 °C to 60 °C over the course of 80 minutes and then from 60 °C to 24 °C over 15 hours. Excessive staples were removed from the folded nanotubes by polyethylene glycol fractionation [102].

To fluorescently label the nanotubes, desired DNA oligonucleotides with complementary sequences to the handles (termed anti-handles, see table S.1.3 for sequences) were added to the monomers (1.2:1 molar ratio between anti-handles and handles) and the hybridization was carried out at 37 °C for two hours. For

dimerization, stoichiometric amounts of the two fluorescently-labeled monomers were mixed and incubated at 16 °C overnight. The final product was loaded on a non-denaturing, 1.5% agarose gel and electrophoresed at 3V/cm in ice-cold 0.5× TBE buffer containing 10 mM $MgCl_2$ for 3.5 hours. It is important not to expose the fluorescently-labeled nanotube dimers to elevated temperature, (>25°C) to avoid cross-contamination because some handles in different fluorescent-labeling zones share the same oligonucleotide sequence (Table S.1.1). The desired band was excised and the nanotube dimers were recovered from the agarose gel using a pellet-pestle method [90]. To prepare an equimolar mixture of 2, 27, 72 or 216 barcodes, all the barcode species involved were folded, labeled and dimerized separately and then mixed at equal molar ratio for agarose gel electrophoresis purification.

To produce the three-point star-like barcode, the DNA nanotube was folded and labeled as described above, and the DNA ring was prepared under similar conditions, with the p3024 scaffold (3,024-base long circular ssDNA derived from pBluescript vector, see table S.1.4 for sequence), staple and $MgCl_2$ concentration changed to 25 nM, 125 nM (each) and 10 mM, respectively. The labeled ring and the nanotube were then purified separately through agarose gel electrophoresis, mixed together at a 1:3 molar ratio and incubated at 37 °C overnight. The final product was purified via gel electrophoresis as described above.

4.9.4 Tagging yeast cells with nano-barcodes

C. albicans yeast cells were suspended in 150 µL of PBS buffer (cells per microliter). 1 µL of 4 mg/mL polyclonal antibody (with or without biotin label) was added to this suspension and allowed to bind to the cell surface for 15 min. The yeasts were then spun in a desktop centrifuge at 4000 rcf for 15 min to form a pellet and washed with 150 µL of PBS buffer twice. The antibody-coated cells were then resuspended in 150 µL of PBS buffer and mixed with 2 µL of 10 mg/mL streptavidin. After incubating for 10 min, the cells were again spun and washed as described above. The cell pellet was then resuspended in 150 µL of PBS buffer containing 10 mM $MgCl_2$ and 0.05% Tween-20 (termed PBS+ buffer here) and treated with 2 µL of 2 nM biotinylated or non-biotinylated GRG nano-barcode. After 30 min incubation on a desktop shaker, the yeast cells were pelleted and washed three times

with 150 µL of PBS+ buffer.

4.9.5 Sample preparation for fluorescence imaging

A piece of coverslip (No. 1.5, 18 18 mm^2, ~0.17 mm thick) and a glass slide (3 1 inch^2, 1 mm thick) were sandwiched together by two strips of double-side tapes to form a flow chamber with inner volume of ~15 µL. First, 20 µL of biotin-labeled bovine albumin (1 mg/mL, dissolved in Buffer A (10 mM Tris-HCl, 100 mM NaCl, 0.05% Tween-20, pH 7.5) was flowed into the chamber and incubated for 2 min. The chamber was then washed using 20 µL of Buffer A twice. 20 µL of streptavidin (0.5 mg/mL, dissolved in Buffer A) was then flown through the chamber and allowed to bind for 2 min. After washing twice with 20 µL of Buffer B (5 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM EDTA, 0.05% Tween-20 pH 8, supplemented with 2.5 mM protocatechuic acid, 10 nM protocatechuate-3,4-dioxygenase and 1 mM Trolox), 20 µL of biotin-labeled nano-barcodes (10 pM in Buffer B) were finally flown into the chamber and incubated for 2 min. The chamber was washed using 20 µL of buffer B three times, sealed with nail polish and mounted on the sample stage of a microscope with the coverslip facing the objective. For DNA-PAINT imaging, samples were prepared as described above. The final imaging buffer solution contained 20 nM ATTO488, Cy3b and ATTO655 labeled imager strands in Buffer B, respectively. For the three-point star-shape barcode and the yeast cells, 5 µL of specimen was directly deposited on the glass slide, covered immediately with a piece of coverslip, sealed with nail polish and incubated for 5 min before being mounted onto the sample stage.

4.9.6 Diffraction-limited fluorescence imaging

Fluorescence imaging was carried out on an inverted Leica DM1600B, microscope (Buffalo Grove, IL) applying an objective-type TIRF configuration using a Leica TIRF illuminator with an oil-immersion objective (HCX PL APO 100/1.47 oil CORR TIRF). Additional lenses were used to achieve a final imaging magnification of ~112 fold, corresponding to a pixel size of ~71 nm. A 488 nm solid state laser (JDS Uniphase, FCD 488-10), a 561 nm solid state laser (LASOS YLK6110) and a 635 nm diode laser (BSR, ChromaLase 635) were used for TIRF excitation. Laser beams were filtered with cleanup filters (zet488x, zet561x, 635x) and coupled into

the microscope objective using a multi-band beamsplitter (zt488, zt561, zt635). Emitted light was spectrally filtered with emission filters (ET525/36, ET600/32, ET705/72) and imaged on an EMCCD camera (Hamamatsu C9100-02, Hamamatsu, Japan). Exposure times were 1 s for 488 nm excitation and 700 ms for 561 and 635 nm excitations. Two consecutive images were taken for the same area of sample, then averaged to yield the final image. The laser beam alignment and imaging processes were automatically controlled by the Leica LAS software. The TIRF images were processed using ImageJ and a custom software for decoding (see the decoding software section).

For epi-fluorescence imaging, the full-spectrum light generated from a mercury lamp was filtered through desired filters (470/40, 560/40 and 645/30 for blue, green and red channel, respectively) to serve as excitation source. The same beam splitter and emission filters were used as in the TIRF mode. The exposure time for all channels was 2 seconds. For the barcode treated yeast sample, an additional bright field image was taken to show the profile of the cells.

4.9.7 DNA-PAINT Imaging

Super-resolution fluorescence imaging using DNA-PAINT was carried out on two different microscope systems. Figure 4 and Figure S.1.14 were acquired on an inverted Zeiss LSM 710 ELYRA PS.1 microscope (Carl Zeiss, Germany) applying an objective-type total internal reflection fluorescence (TIRF) configuration using an oil immersion objective (100x alpha Plan Apo, NA 1.46, Oil, Zeiss). An additional 1.6 magnification was used to obtain a final imaging magnification of ~160 fold, corresponding to a pixel size of 100 nm. Samples were imaged using custom made microscopic chambers. Three lasers (λ=488 nm, 100 mW, λ=561 nm, 100 mW and λ=642 nm, 100 mW) were used for TIRF excitation. The laser beam was coupled into the microscope objective using a multi-band beamsplitter. Emitted light was spectrally filtered with an emission filter (BP 495-575/LP 750 for ATTO488 imaging, BP 570-65-/LP 750 for Cy3b imaging and LP 655 for ATTO655 imaging) and imaged on an EMCCD camera (Andor iXon 3, Andor Technologies, North Ireland). 5000 frames were recorded at a frame rate of 50 Hz for each color channel sequentially.

Figure S.1.13 was acquired on an inverted Nikon Ti-E microscope (Nikon

Instruments, Melville, NY) with the Perfect Focus System, applying an objective-type TIRF configuration using a Nikon TIRF illuminator with an oil-immersion objective (100 Plan Apo, NA 1.49, Oil, Nikon). An additional 1.5x magnification was used to obtain a final imaging magnification of ~150 fold, corresponding to a pixel size of 107 nm. Three lasers (λ=491 nm, 100 mW, λ=561 nm, 75 mW and λ=640 nm, 30 mW, BSR) were used for TIRF excitation. The laser beam was filtered with a cleanup filter (488/20 for ATTO488 imaging, 561/20 for Cy3b imaging and 642/20 for ATTO655 imaging, all Chroma Technologies) and coupled into the microscope objective using a multi-band beamsplitter (zt405/488-491/561/638rpc, Chroma Technologies). Emitted light was spectrally filtered with an emission filter (525/50 for ATTO488 imaging, 600/50 for Cy3b imaging and 700/75 for ATTO655 imaging, all Chroma Technologies) and imaged on an EMCCD camera (Hamamatsu ImagEM C-9100-13, Hamamatsu, Japan). 5000 frames were recorded at a frame rate of 10 Hz for each color channel sequentially.

Super-resolution images were reconstructed using spot-finding and two-dimensional Gaussian fitting algorithms programmed in LabVIEW [144]. A simplified version of this software is available for download at `www.e14.ph.tum.de`. Images were drift corrected, and the different channels were aligned using DNA origami drift markers with binding sites for all three imager strands (see Figure S.1.1d). The origami structure is based on Rothemund's rectangular-shaped origami [135]. The black staple strands in the strand diagram were randomly extended with two out of three docking/handle strands for DNA-PAINT imaging. The resulting structure displays ~80 binding sites on average for the red, green and blue imager strands, respectively. This high density of binding sites ensure that at any given time one or more image strands will be in the bound state. This makes the structures ideal for systematic drift correction in each channel. Because the structures are imaged in each channel at the same positions, they can be used to correct offsets and align the red, green and blue channel accordingly.

4.9.8 Automated characterization of the geometry of BRG and GRG barcodes

For the purpose of assessing barcode geometry, we chose BRG and GRG barcodes as examples. The software first determined the centroids of all visible spots in

the diffraction-limited TIRF images based on their fluorescence intensities. The centroid of a spot was defined as the brightest central pixel in a 3-pixel by 3-pixel region. The software processed blue, green and red fluorescence channels independently to identify all the centroids in each channel (Figure S.1.4). It then superimposed the three channels and found the correct barcodes based on the design blueprint. The distances between the neighboring fluorescently-labeled zones were calculated as the distances between the centroids of the adjacent spots on a barcode. The barcode bending angle a was calculated using the cosine formula as shown below, where a and b are the distances between neighboring centroids and c is the distance between the farthest two centroids within a barcode. The measured parameters are listed in Table S.1.5.

4.9.9 Decoding software for diffraction-limited TIRF images

Custom MATLAB scripts were used to decode barcodes from diffraction-limited TIRF images. The software can run in either a supervised or an unsupervised mode. In the supervised mode, each decoded barcode is presented to the user to accept or reject. In unsupervised mode, the software decodes all barcodes without any user input.

The software operates in two steps: First, the software identifies the location and position of barcodes. Second, the software actually decodes the barcode (See Figure 4.7). For step one, the software preprocesses the three-channel TIRF image containing the barcodes (one channel each for fluorescence in the red, green and blue). The image is spatially high-pass filtered to remove slowly spatially varying background illumination. Each color channel is then thresholded independently so that only pixels corresponding to putative barcodes remain. The three-channel thresholded image is then merged to generate a single-channel binary image. The software identifies shapes from the binary image that match the length and width criteria of a barcode. The position and orientation of these shapes are recorded. In step two, the software reads each barcode identified in step one. For each barcode, the software extracts the corresponding region of the original three-channel image. This constitutes the observed barcode. The software then performs a matched filter to compare the observed barcode against a library of all possible reference barcodes. For each reference, three simultaneous correlations are performed, one

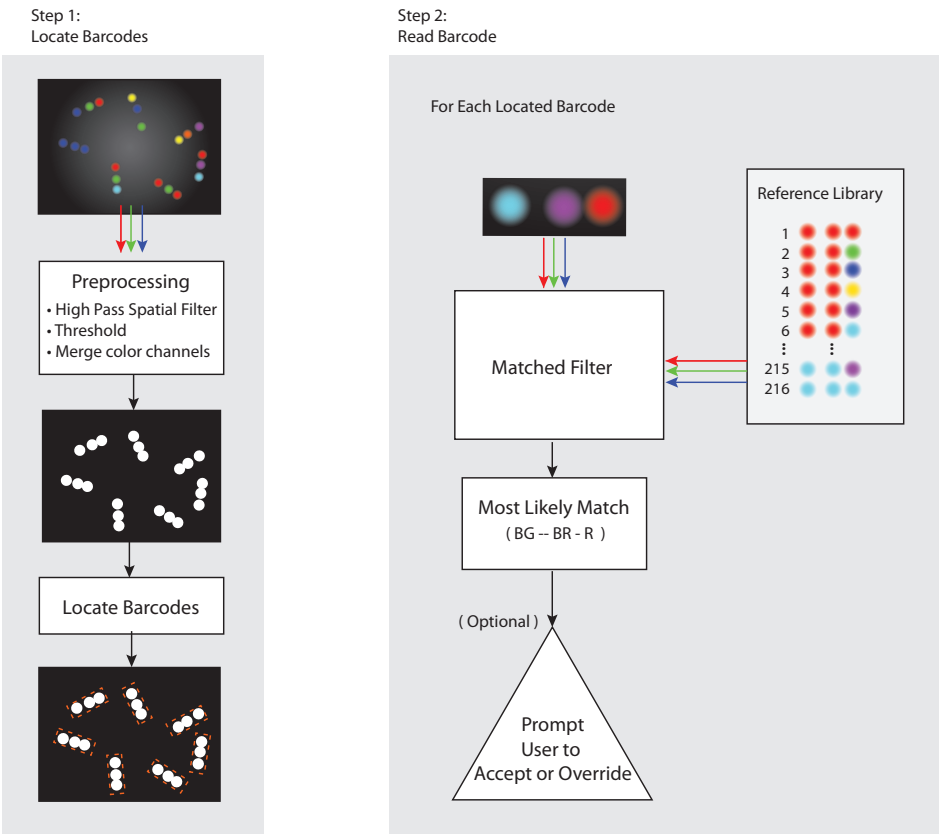


Figure 4.7: Barcode analysis software box diagram. The MATLAB software operates in two steps: In the first step the software identifes the location and position of barcodes. In the second step the software decodes the barcode using a matched filter.

in each color channel, while keeping all channels aligned. To reduce computations, these correlations are actually performed on one-dimensional projections of the two-dimensional image of the observed barcode in each channel, so that each fluorescent spot appears as a peak in a one-dimensional trace. The observed barcode is assigned the identity of the reference barcode with the highest combined correlation across all three color channels.

In supervised mode, the computer will present the user with the image of the barcode and its most likely identity. The user can either accept the computers proposal or manually enter the correct identity of the barcode. In cases where the

software erroneously decodes misshapen or malformed barcodes, the user can reject the barcode entirely. The software finally outputs the counts of all barcode species as well as the coordinates and identity of each individual barcode. In supervised mode, the computer also registers the number of occasions when the user approves the its proposal. For the 72-barcode pool, this agreement rate is approximately 80%.

4.10 Acknowledgements

We thank Dr. Christian Steinhauer and Sebastian Piet Laurien for help with super-resolution software development. We thank Dr. Bernhard Goetze and the Harvard Center for Biological Imaging as well as the Nikon Imaging Center at Harvard Medical School for the use of their microscopes. We thank Dr. Shawn M. Douglas for providing TEM images used in Supplementary Figure S.1.5. This work was supported by grants from an NIH Director's New Innovator Award and a Wyss Institute for Biologically Inspired Engineering Faculty Award to W.M.S. and an NIH Director's New Innovator Award, an NSF Faculty Early Development Award, an ONR Young Investigator Program Award, and Wyss Institute for Biologically Engineering Faculty Startup Fund to P.Y. R.J. acknowledges support from the Alexander von Humboldt-Foundation through a Feodor Lynen fellowship.

4.11 Manuscript Information

4.11.1 Submitted for Publication As

A version of this chapter has been submitted for publication to the journal *Nature Methods* as [101],

Chenxiang Lin, Ralf Jungmann, Andrew M Leifer, Chao Li, Daniel Levner, William M Shih, and Peng Yin. Sub-micrometer geometrically encoded fluorescent barcodes Self-Assembled from DNA. Submitted to Nature Nanotechnology, October 2011.

4.11.2 Author Contributions

The authors for this work were: Chenxiang Lin, Ralf Jungmann Andrew M. Leifer, Chao Li, Daniel Levner, William M. Shih, and Peng Yin.

Andrew M. Leifer designed and wrote the software to analyze TIRF images and identify and decode the barcodes. He wrote portions of the manuscript and generated Figure 4.7. This work lays the foundation for more detailed analysis performed in Chapter 5.

C. Lin conceived the project, designed and conducted the majority of the experiments, analyzed the data and prepared the majority of the draft. R.J. conceived the super resolution barcode study, designed and conducted experiments for this study, analyzed the data, and prepared the draft. C. Li wrote the C script for the barcode geometry characterization. D.L. (with C. Lin) performed the yeast tagging experiment. W.M.S. conceived the project, discussed the results, and prepared the draft. P.Y. conceived, designed, and supervised the study, interpreted the data, and prepared the draft. All authors reviewed and approved the draft.

4.12 Supplementary Material

All Supplementary Tables and Figures from this chapter are located on the accompanying CD-ROM in the file: `Chapter4Supplementary.pdf`.

In summary, there are no small problems. Problems that appear small are large problems that are not understood.

Ramón y Cajal, [20]

5

Automated DNA origami barcode detection using Bayseian multiple hypothesis testing

5.1 Introduction

The development of next generation microarray and on-chip technologies will depend in part upon advancements in fluorescent reporter molecules and improvements in their detection.

Currently-accepted microarray techniques use at most two spectrally-distinct fluorophores to identify their targets. Recently, however, there has been an explosion in fluorescent reporter approaches that enable an observer to uniquely identify one out of many hundreds of distinct fluorescent encoded reporters. Researchers have explored both combining spectrally-distinct fluorophores in close proximity to create gradations of colors [51, 63, 96, 99, 105, 112, 166], and spatially separating fluorophores to create geometric optical barcodes [17, 36, 55, 59,

95, 120, 128, 165]. In particular, fluorescently-encoded DNA origami [41, 135] based nanorods (Chapter 4) provide a promising avenue for DNA based microarray techniques. Instead of identifying microarray targets largely by location, this raises the distinct possibility of also unambiguously identifying targets by fluorescent encoding.

However, programmatically identifying images of these fluorescently-encoded barcodes, presents unique computational challenges that have not yet been explored in the literature. Here, we describe a mathematical formalism that allows us to use computer vision to automatically optimally detect DNA origami-based nanorod barcodes. We take a bayesian inference approach and demonstrate that our implementation allows for >95 % correct barcode identification when tested against barcodes pre-selected to be well formed and evenly labeled.

5.2 System and Methods

5.2.1 DNA origami nanorod barcodes

We chose to image DNA origami nanorod barcodes (see Chapter 4). These nanorods consist of mechanically rigid six-helix bundles of DNA with radius 5 nm and length 720 nm decorated with 12 molecules of up to two spectrally-distinct fluorophores at each of three distinct locations along the rod, as illustrated in Figure 5.1. The three fluorescing sites are unevenly spaced, with a large gap of ~450 nm and a small gap of ~270 nm between the fluorescing sites on each rod, giving the barcodes a visible anisotropy. When using combinations of red, green and blue fluorophores, (Cy5, Cy3 and Alexa Fluor 488, respectively) this allows for a dictionary of six fluorescing symbols at each of the three barcode locations: {Red, Green, Blue, RedGreen, RedBlue, GreenBlue}, for a total of $6^3 = 216$ unambiguous species of barcode.

We chose to work with this particular form of fluorescently-encoded barcodes as opposed to others because this form offer the largest published number of distinct barcode species to date, has the advantage of being massively self-assembled in parallel and creates barcodes of suitably small size so as to be useful for future microarray applications.

The barcodes are, in fact, so small that the distance between fluorescing sites is

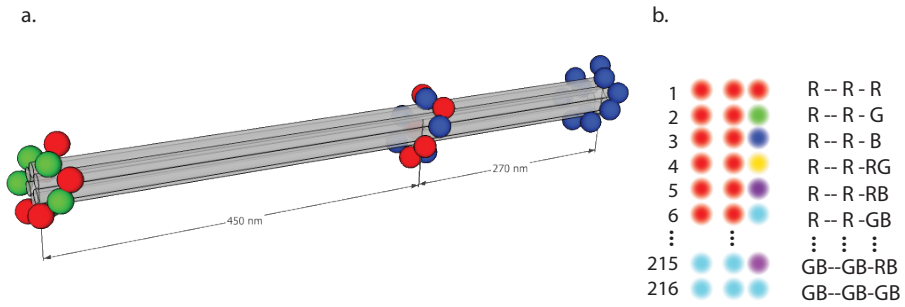


Figure 5.1: Cartoon of of DNA origami nanorod barcodes (not to scale) . **(a)** The barcode is a composed of a six helix-bundle of DNA, 5 nm in diameter and 720nm long, decorated at three fluorescing sites with 12 molecules of up to two spectrally distinct fluorophores per site. The barcode has an inherent anisotropy due to the uneven spacing of three fluorescing sites (large gap on the left, smaller gap on the right). The barcode shown is of species RedGreen–RedBlue-Blue. **(b)** The barcodes use a combination of red (R), green (G), and blue (B) fluorescent molecules, providing a dictionary of six fluorescing symbols at each of the three barcode locations {R, G, B, RG, RB, GB}, for a total of $6^3 = 216$ distinct barcode species. In composite images RG appears yellow, RB appears purple and GB appears turquoise.

often smaller than a wavelength of light. As a result, when the barcodes are imaged on a flat surface, the distinct flourescent sites appear to blend into one another (see Figure 5.2). This becomes one of the primary challenges facing any computer vision barcode identification software–namely to decode the barcode even when there is overlap between individual difraction-limited fluorescent spots.

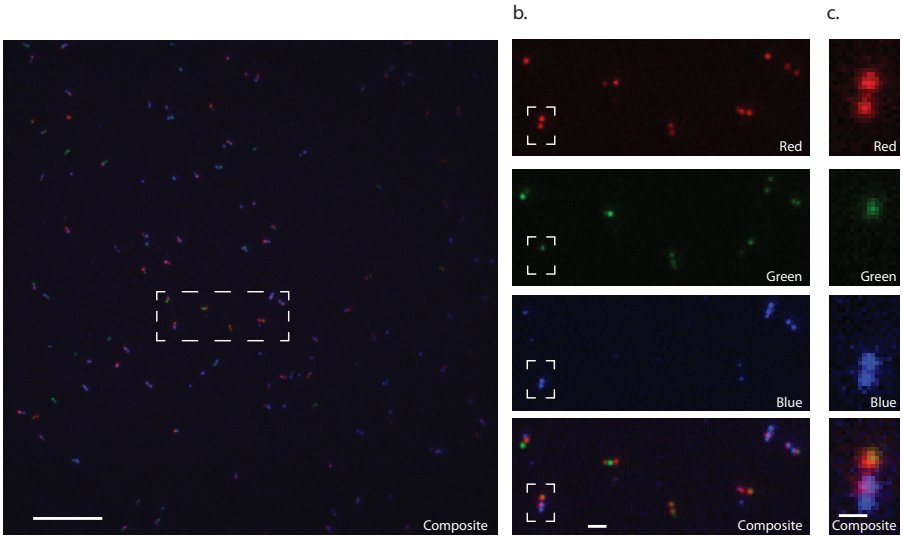


Figure 5.2: Fluorescent images of DNA origami nanorod barcodes. **(a)** Composite image. Scale bar is 100 µm. **(b)** Inset of dashed region in **a**. Scale bar is 10 µm. **(c)** A single barcode of species RedGreen–RedBlue-Blue. (Inset of dashed region in **c**.) Scale bar is 500 nm.

5.2.2 Sample Preparation

Barcodes were prepared using the same technique described in Chapter 4. For imaging, the biotin-labeled barcodes were immobilized on a glass coverslide functionalized with biotin-streptavanin linkers.

5.2.3 Imaging

Barcodes were imaged on a Leica DM1600 microscope using a 100x oil-immersion objective, under total internal reflection fluorescence (TIRF). For details, see

Chapter 4. Breifly, the red, green and blue channel images of the barcodes were recorded using 635 nm, 561 nm and 488 nm laser light, respectively, from diode pumped solid state lasers (BSR, ChromaLase 635; LASOS YLK6110; and JDS Uniphase, FCD 488-10) Emitted light was spectrally filtered with emission filters (ET525/36, ET600/32, ET705/72) and imaged onto an electron-multiplying CCD camera (Hamamatsu C9100-02, Hamamatsu, Japan). Table 5.1 lists exposure times and other image information.

Image Properties

Parameter	Value
Scale	1 px $\approx$ 71.4 nm
Size	1,000 px $\times$ 1,000 px
Bit Depth	14 bits
Exposure	700 ms

Table 5.1: Image properties.

5.2.4 Task at Hand

To be useful for microarray or other diagnostic applications, it is important for any computer vision software to be able to automatically detect and decode large quantities of barcodes robustly and accurately with little or no user input. Here, we restrict ourselves to immobilized barcodes laying flat on a glass slide, but we demand that our software solution tolerate likely real-world complexities including the presence of inhomogeneous spatial illumination, background fluorescence and modest flexing of the barcodes. Even so, we demand that the barcodes be evenly fluorescently labeled, so that the same number of fluorophores are present on each fluorescing spot. While only a modest percentage of barcodes manufactured using current techniques meet this criteria, DNA barcode technology is in its infancy, and we are confident that manufacturing quality and robustness will improve dramatically in coming years. For now, we pre-screen for barcodes that have even fluorescence.

In the following sections, we develop an optimal mathematical framework and software implementation for identifying the location and orientation of barcodes

and then decoding them.

5.3 Theory

To automatically detect and decode a barcode from its three-channel image, the software must systematically scan through the image, detect the location of each barcode and decide which of the 216 reference barcode species best fits the observed barcode. Assigning the best fitting reference barcode is performed by the Bayesian multiple hypothesis tests to be described shortly. Before performing the tests, the raw image data must be conditioned to remove DC offsets and scaling introduced by the optical system and by the camera's charge-coupled-device (CCD) detector.

5.3.1 Conditioning the Raw Image

The first step in processing the image is removing the offsets and correcting the scaling, both of which vary across the image as a consequence of the illumination and optical system properties (See Figure 5.3a). We begin by describing the signals and their sources.

The voltage $V(\mathbf{r})$ at a pixel $\mathbf{r}$ in the detector is the sum of a term that is proportional to the number of incident photons, a DC offset $V_o(\mathbf{r})$ that can slowly vary from pixel to pixel, and thermal noise v_t that we assume is Gaussian and zero-mean,

$$V(\mathbf{r}) = V_o(\mathbf{r}) + c \cdot \text{Photons}(\mathbf{r}) + v_t, \tag{5.1}$$

where c is the conversion constant that relates the number of photons to voltage.

In our system, laser light at one wavelength is incident on the sample and stimulates a fluorophore to emit photons at another wavelength that passes through a narrowband filter to reach the detector. Photons that reach the detector must have been emitted either from fluorophores located in the barcodes themselves f or from contaminants and impurities in the background media b, which consists of a glass slide and buffer. In both cases, we model the number of photons that are emitted as proportional to the number of incident photons,

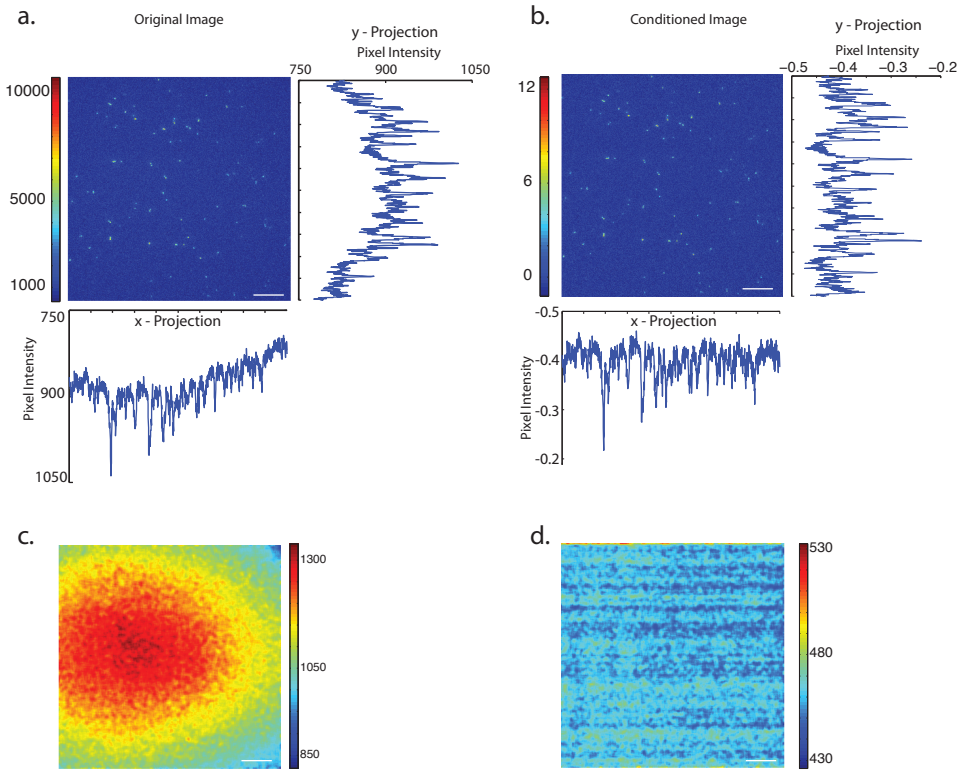


Figure 5.3: The raw image must be conditioned to remove offsets and to correct scaling. For all images, scalebar is 10 μm.(**a**) The raw red channel image (1000 px by 1000 px) of DNA origami barcodes is shown along with its projections onto the x- and y-axes of the image. The raw image has a large offset and the background varies slowly. (**b**) The conditioned red channel image is shown with projections onto the x and y axes. The background of the conditioned image has a mean that is constant and nearly zero. **(c)** Image of only the red channel background, at magnified scale. This image, which has been smoothed by convolution with a Gaussian of width $\sigma =$ 10 px, was taken under the same conditions as **a**. It contains slide and buffer, without any barcodes. This image is used to correct for uneven illumination, visible as the glowing spot towards the left. **(d)** Dark red channel image taken with neither slide, nor buffer, nor illumination, is used to correct for DC camera offset. This image has also been smoothed by convolution with a Gaussian of width $\sigma = 10$ px.

$$f(\mathbf{r}) = \gamma_{\text{ex}}(\mathbf{r}) q_f(\mathbf{r}) \tag{5.2}$$

$$b(\mathbf{r}) = \gamma_{\text{ex}}(\mathbf{r}) q_b. \tag{5.3}$$

Here, $\gamma_{\text{ex}}(\mathbf{r})$ is the slowly varying position-dependent intensity of the excitation laser; $q_f(\mathbf{r})$ describes the position-dependent fluorescent efficiency of the barcodes independent of illumination; and q_b is the background fluorescent efficiency of the media which is assumed to be independent of position and illumination. The term q_f is itself composed of light from fluorophores at positions $\mathbf{r}_i$, each modulated by the point spread function (PSF) h of the optical system

$$q_f(\mathbf{r}) = \sum_j a_j \delta(\mathbf{r} - \mathbf{r}_j) * h(\mathbf{r}). \tag{5.4}$$

Here a_j is a factor that accounts for intensity variations due to inconsistent fluorescent labeling during barcode assembly, and the star denotes the convolution operation.

The contributions of the terms in Eq. 5.1 can be evaluated by acquiring auxiliary images. The offset and thermal detector noise are present in a dark image taken when the laser is off, for example. The background media noise and spatially varying illumination become visible in an image taken with the laser on and with media present but with no barcode samples. The final image, of course, has all of these plus DNA barcodes in the media. The voltages recorded in these cases are

$$V_{\text{dark}}(\mathbf{r}) = V_{\text{o}}(\mathbf{r}) + v_t \tag{5.5}$$

$$V_{\text{back}}(\mathbf{r}) = V_{\text{o}}(\mathbf{r}) + c \cdot b(\mathbf{r}) + v_t \tag{5.6}$$

$$V_{\text{fluor}}(\mathbf{r}) = V_{\text{o}}(\mathbf{r}) + c \cdot \left[f(\mathbf{r}) + b(\mathbf{r}) \right] + v_t \tag{5.7}$$

corresponding to Figure 5.3d, c and a, respectively.

The offset and illumination intensity vary over the length of the image but are constant over the length of a barcode. To characterize the variations while removing rapid noise fluctuations, the dark and background images are smoothed over a length scale large compared to a barcode but small compared to the entire image. These smoothed images are then subtracted and divided to remove the DC offset

and compensate for the spatially varying illumination, respectively, producing a unitless conditioned image (Figure 5.3b),

$$U(\mathbf{r}) = \frac{V_{\text{fluor}}(\mathbf{r}) - \widetilde{V}_{\text{back}}(\mathbf{r})}{\widetilde{V}_{\text{back}}(\mathbf{r}) - \widetilde{V}_{\text{dark}}(\mathbf{r})}, \tag{5.8}$$

where $\widetilde{\ }$ denotes smoothing. Substituting Eq. 5.2 gives

$$U(\mathbf{r}) = \frac{c \cdot \left[f(\mathbf{r}) + b(\mathbf{r}) - \widetilde{b}(\mathbf{r})\right] + v_t}{c \cdot \widetilde{b}(\mathbf{r})} \tag{5.9}$$

Note that $\widetilde{V}_{\text{o}}(\mathbf{r}) = V_{\text{o}}(\mathbf{r})$ and $\widetilde{\gamma}_{\text{ex}}(\mathbf{r}) = \gamma_{\text{ex}}(\mathbf{r})$ since both are assumed to be slowly varying. Substituting further, and recalling that the means of the noise terms are zero, gives

$$U(\mathbf{r}) = \frac{c \cdot \gamma_{\text{ex}}(\mathbf{r}) q_f(\mathbf{r}) + v_t}{c\gamma_{\text{ex}}(\mathbf{r}) q_b} \tag{5.10}$$

or

$$U(\mathbf{r}) = \frac{q_f(\mathbf{r})}{q_b} + \frac{v_t}{c\gamma_{\text{ex}}(\mathbf{r}) q_b}. \tag{5.11}$$

The first term is the fundamental fluorophore distribution; the second term is thermal noise scaled by the local intensity of illumination; while the factor q_b is a constant. This conditioned thermal noise term is zero-mean and Gaussian to reasonable approximation, as will be verified later.

Now that the spatially-varying offset and illumination have been removed, we turn to the problem of barcode identification. Since performing multiple-hypothesis testing on a two-dimensional dataset is computationally intensive, we start with a simple one-dimensional identification problem.

5.3.2 One dimension, single channel

In the one-dimensional problem, we assume that there are K possible one-dimensional reference signals in a system with zero-mean additive Gaussian noise n, such that for any reference m_k, the observed signal $u = m_k + n$. Collecting each signal into a data vector gives

$$\mathbf{u} = \mathbf{m} + \mathbf{n}, \tag{5.12}$$

Given a noisy observation, $\mathbf{u}$, we seek the most likely corresponding reference, or model signal $\mathbf{m}_k$.

The inverse probability that the kth reference signal $\mathbf{m}_k$ is present when the data vector $\mathbf{u}$ is observed is given by Bayes' theorem

$$p(\mathbf{m}_k|\mathbf{u}) = \frac{p(\mathbf{u}|\mathbf{m}_k)p(\mathbf{m}_k)}{p(\mathbf{u})}, \tag{5.13}$$

where $p(\mathbf{u}|\mathbf{m}_k)$ is the forward probability or likelihood of seeing data $\mathbf{u}$ given the presence of model $\mathbf{m}_k$, $p(\mathbf{m}_k)$ is the prior probability that model $\mathbf{m}_k$ is present, and $p(\mathbf{u})$ is a normalization constant [77]. We seek to compute the inverse probability for each possible reference $\mathbf{m}_k$, $k = 1, 2, 3 \ldots K$ and select the largest, which is the one most likely to fit the observed data.

While testing a particular observed signal $\mathbf{u}$ against all models, the denominator is constant and independent of k. Maximizing $p(\mathbf{m}_k|\mathbf{u})$ is therefore equivalent to finding the model k that maximizes the product $p(\mathbf{u}|\mathbf{m}_k)p(\mathbf{m}_k)$. In the present case, where all barcodes are equally probable, $p(\mathbf{m}_k) = 1/K$ and the search simplifies to finding the maximum likelihood

$$L_k = p(\mathbf{u}|\mathbf{m}_k). \tag{5.14}$$

This is termed the maximum likelihood solution.

The probability $p(\mathbf{u}|\mathbf{m}_k)$ of observing $\mathbf{u}$ given the reference $\mathbf{m}_k$ is simply the probability that the difference $\mathbf{u} - \mathbf{m}_k$ between the signal and the reference is generated by noise. This is given by the multivariate probability density function for the random Gaussian variable $(\mathbf{u} - \mathbf{m}_k)$,

$$L_k = p(\mathbf{u}|\mathbf{m}_k) = \frac{1}{\sqrt{(2\pi)^N \det ||\mathbf{R}_n||}} \exp\left[-\frac{1}{2}(\mathbf{u} - \mathbf{m}_k)^T \mathbf{R}_n^{-1}(\mathbf{u} - \mathbf{m}_k)\right] \tag{5.15}$$

where N is the length of vector $\mathbf{u}$; $\mathbf{R}_n$ is the noise covariance matrix; $\det ||\cdot||$ is the matrix determinant; and T denotes the transpose operation [72, 157].

For the case of identical independently distributed (IID) Gaussian noise,

$$\mathbf{R}_n = \sigma^2 \mathbf{I} \tag{5.16}$$

where $\mathbf{I}$ is the identity matrix and σ^2 is the variance of the noise. The probability thus reduces further to

$$L_k = \frac{1}{\sigma^N \sqrt{(2\pi)^N}} \exp\left[-\frac{(\mathbf{u}-\mathbf{m}_k)^T(\mathbf{u}-\mathbf{m}_k)}{2N\sigma^2}\right]. \tag{5.17}$$

The exponent is simply the energy contained in the sequence $\mathbf{u}-\mathbf{m}_k$ divided by the energy of a sequence of noise of length N and variance σ^2. We recognize parallels to the Maxwell-Boltzmann distribution of statistical mechanics [131], whereby high-energy sequences (in this case ones where the model is a poor match to the observed signal) are unlikely to be generated by random noise.

As mentioned, we find the $\mathbf{m}_k$, or rather just the index k, that produces the maximum likelihood in equation 5.17, that is,

$$\max_k \{L_k\} = \max_k \left\{ \frac{1}{\sigma^N \sqrt{(2\pi)^N}} \exp\left[-\frac{(\mathbf{u}-\mathbf{m}_k)^T(\mathbf{u}-\mathbf{m}_k)}{2N\sigma^2}\right] \right\}. \tag{5.18}$$

This equation leads directly to implementation with a classic bank of matched filters that compares observed and reference waveforms [72, 157].

This equation has another interpretation as well. Since σ and N are independent of k, and since the exponential is a monotonically increasing function of its argument, maximizing the likelihood across k is equivalent to minimizing the energy of the difference between the observation and the reference across k

$$\max_k \{p(\mathbf{m}_k|\mathbf{u})\} = \min_k \left\{(\mathbf{u}-\mathbf{m}_k)^T(\mathbf{u}-\mathbf{m}_k)\right\}. \tag{5.19}$$

Note that the Bayesian hypothesis test is equivalent in this case to finding the k that minimizes the mean squared error between observation and reference. This is clear by rewriting this expression explicitly in terms of a sum over the components in the waveform vectors,

$$\max_k \{p(\mathbf{m}_k|\mathbf{u})\} = \min_k \left\{ \sum_{j=1}^{N} (u_j - m_{k,j})^2 \right\}. \tag{5.20}$$

Thus, for the one-dimensional barcode case, the Bayesian hypothesis test is equivalent to both classic matched filter and least-mean-squared-error (LMSE) detec-

tion strategies. All are strictly optimum strategies for the problem posed here.

5.3.3 One dimension, multiple channels

We now consider a system where each reference $\mathbf{m}_k$ consists of three one-dimensional vectors $\mathbf{m}_{k1}$, $\mathbf{m}_{k2}$ and $\mathbf{m}_{k3}$, each of length N. Each will correspond to the portion of the reference contained in a different channel so that there are still K references. Analogously, the signal $\mathbf{u}$ also consists of three one-dimensional observations $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$, each in its own channel and each channel has zero-mean additive IID Gaussian noise with variances σ_1^2, σ_2^2 and σ_3^2, respectively. As we will see later, these channels can be thought of as the red, green and blue channels in a one-dimensional composite color image.

Because the three channels are independent, the likelihood that signal $\mathbf{u}$ is observed when $\mathbf{m}_k$ is present simply factors into the product of the individual likelihoods. Namely, it is the likelihood that $\mathbf{u}_1$ is observed when $\mathbf{m}_{k1}$ is present times the likelihood that $\mathbf{u}_2$ is observed when $\mathbf{m}_{k2}$ is present times the likelihood that $\mathbf{u}_3$ is observed when $\mathbf{m}_{k3}$ is present [77],

$$L_k = p(\mathbf{u}|\mathbf{m}_k) = p(\mathbf{u}_1|\mathbf{m}_{k1})p(\mathbf{u}_2|\mathbf{m}_{k2})p(\mathbf{u}_3|\mathbf{m}_{k3}). \tag{5.21}$$

Substituting Eq. 5.17 gives

$$L_k = \frac{1}{(\sigma_1\sigma_2\sigma_3)^N(2\pi)^{3N/2}} \exp\left[-\frac{(\mathbf{u}-\mathbf{m}_{k1})^T(\mathbf{u}-\mathbf{m}_{k1})}{2N\sigma_1^2} - \frac{(\mathbf{u}-\mathbf{m}_{k2})^T(\mathbf{u}-\mathbf{m}_{k2})}{2N\sigma_2^2} - \frac{(\mathbf{u}-\mathbf{m}_{k3})^T(\mathbf{u}-\mathbf{m}_{k3})}{2N\sigma_3^2} \right]. \tag{5.22}$$

As before, we can factor out and ignore terms that have no k dependence. Thus, the maximum likelihood model k corresponds to finding k that minimizes the

quantity on the right below:

$$\max_k \{L_k\} = \min_k \left\{ \frac{(\mathbf{u}_1 - \mathbf{m}_{k1})^T (u_1 - \mathbf{m}_{k1})}{\sigma_1^2} + \frac{(\mathbf{u}_2 - \mathbf{m}_{k2})^T (\mathbf{u}_2 - \mathbf{m}_{k2})}{\sigma_2^2} + \frac{(\mathbf{u}_3 - \mathbf{m}_{k3})^T (\mathbf{u}_3 - \mathbf{m}_{k3})}{\sigma_3^2} \right\}. \quad (5.23)$$

As before, this maximum likelihood formalism leads to a classic matched filter bank and is the optimal solution to finding the identity of the barcode.

5.4 Implementation

To identify and decode our barcodes, we start with three conditioned images, $U_1(\mathbf{r}), U_2(\mathbf{r}), U_3(\mathbf{r})$, one each for channels red, green, and blue, respectively. Before we can begin to employ the solution in Eq. 5.23, we must locate the barcodes within our image, normalize the intensities of the barcodes across the three channels and project the barcodes down to a single dimension (See Figure 5.4). Each of these steps requires multiple sequential image manipulations, which are described in detail in the next three sections. At the conclusion of the first, Section 5.4.4, we will have one-dimensional 3-channel projections of our barcodes. Section 5.4.5 describes how the reference signals are generated. Finally, in Section 5.4.6, the barcodes are decoded by comparing the observed signals with the generated references using Eq. 5.23.

5.4.1 Locating the Barcodes

Before we can decode the barcodes, we first must locate them in our image. We locate the barcodes by thresholding away background noise and studying the morphology of the remaining bright image features. Finally, we apply selection criteria to select out only image features that match the shape of a barcode (see Figure 5.5).

To threshold, we generate a smoothed image $\tilde{U}_i(\mathbf{r})$ for each channel by convolving U_i with a Gaussian whose width is narrower than that of the point spread func-

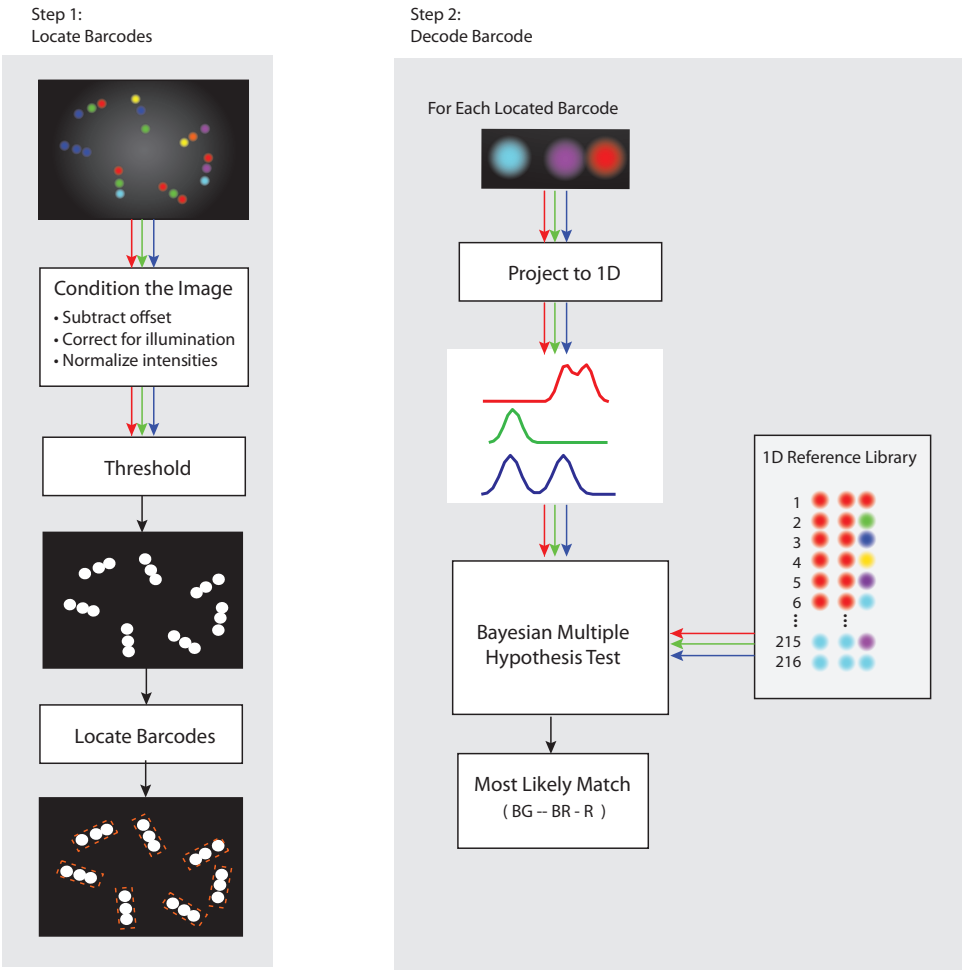


Figure 5.4: Block diagram of the implementation. There are multiple image processing steps to prepare the image and locate the barcode before it can be decoded using a Bayesian multiple hypothesis test.

tion in Eq. 5.4. See Figure 5.5b. The user manually selects a threshold t_i for each of the three channels of the smoothed conditioned image to generate a binary mask,

$$B_i(\boldsymbol{r}) = \begin{cases} 1 & \text{if } \tilde{U}_i(\boldsymbol{r}) \geq t_i, \\ 0 & \text{if } \tilde{U}_i(\boldsymbol{r}) < t_i, \end{cases} \qquad \text{for } i = 1, 2, 3 \tag{5.24}$$

as in Figure 5.5c. Noise below the threshold is masked out. Recall that unlike the original images, the conditioned images U_i (Figure 5.5a) have homogeneous zero-mean backgrounds, so a single threshold per channel is sufficient to capture all barcodes. The three channels of binary masks are merged,

$$B_\cup = B_1 \cup B_2 \cup B_3, \tag{5.25}$$

as in Figure 5.5d. $B_\cup$ consists of many small discontiguous islands, here called blobs, against a masked-away zero background. We can locate barcodes by examining the patterns of blobs.

Some barcodes lie entirely within one blob, while others are represented by two or more nearby but distinct blobs–so we must first group together adjacent but distinct blobs. To do this, we dilate the binary mask using standard image processing routines [114] so that closely spaced blobs run together into single larger blobs. We can label the larger blobs and use them to group together their constituent blobs, so that if a barcode is made up of two nearby blobs in $B_\cup$, they both will be labeled together as one blob ensemble.

Morphological Criteria for Barcode

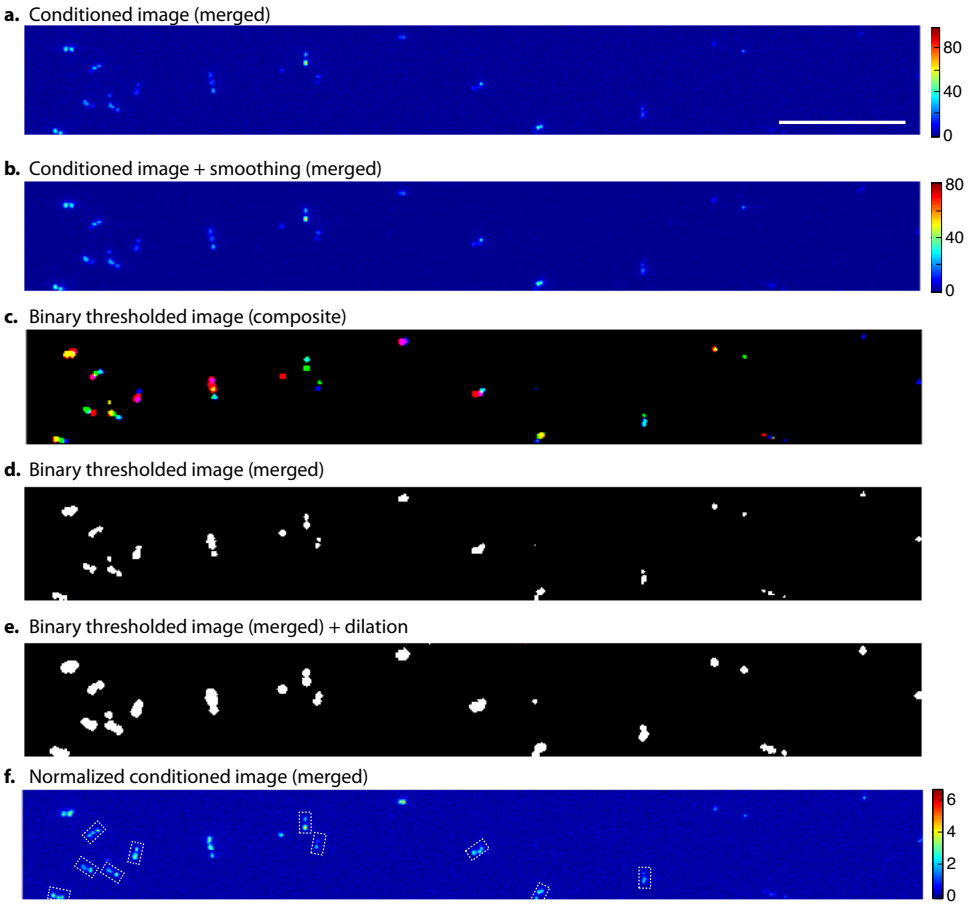
Parameter	Value
Min Area	24 px^2
Max Area	134 px^2
Min Length of Major Axis	14 px^2
Min Eccentricity	0.75

Table 5.2: Morphological criteria for barcode. Scale: 1 px $\approx$ 71.4 nm.

Each blob ensemble is examined to determine if it represents a barcode. To qualify as a barcode, the area of a blob ensemble must lie in an acceptable range, the major and minor axis of the ellipse that fits its second moment must lie in a range

Figure 5.5 ***(following page)*****:** Image processing steps used to locate barcodes and to normalize the intensity across the three channels. **(a)** The conditioned image of the barcodes is shown. The three color channels have been merged into a single channel. Color indicates intensity and scale bar is 10 μm long. **(b)** Each channel of the conditioned image in **a** has been smoothed by convolution with a Gaussian of width less than that of the average peak width. This decreases noise for thresholding. The three smoothed color channels are shown here merged together, where color indicates intensity. **(c)** Each channel of the smoothed conditioned image in **b** is thresholded according to a channel-specific user-specified threshold. The three binary images, one each for channel Red, Green and Blue, are shown in a composite image. By inspecting the thresholded regions in each color channel, the software can measure the mean peak height for each channel. **(d)** The union of the three binary images in **c** is shown as a single binary image. Contiguous white patches are called blobs. **(e)** The binary image in **d** is dilated. This image is used by the software to group nearby blobs in image **d** together into blob ensembles. The area and eccentricity of each blob ensemble is inspected, and those that match the criteria of a barcode are flagged as such. Their location and orientation is recorded. **(f)** The conditioned image normalized by mean peak height in each channel shows uniform fluorescing intensity compared to **a**. On average the height of peaks in each channel is unity. Here the the three channels are shown merged as one, where color indicates intensity. Dotted lines indicate the location of blob ensembles from **d** that match the criteria for barcodes.

Figure 5.5: (continued)



consistent with that of a barcode and the ellipse must be sufficiently elongated or eccentric. Table 5.2 lists the pertinent numerical criteria.

The software also records the centroid and the angle of the barcode major axis, which will be used in Section 5.4.4 to rotate the barcode and project it to one dimension.

5.4.2 Normalizing Intensity of Each Channel

We continue conditioning the raw image in this section. The average barcode intensity can vary across color channels since it depends on channel-specific characteristics like the fluorophore, illumination wavelength, illumination intensity, and the width of the narrowband filter. These are normalized across channels by dividing the observed intensity in each channel by the mean peak height in that channel, $\langle g_i \rangle$, where $i = 1, 2, 3$ denotes the channel number.

To measure mean peak height, the software first identifies blobs in the binary mask B_i that have an eccentricity and an area consistent with that of a single peak. It then finds the brightest pixel in each blob that matches the criteria described in Table 5.3.

Criteria for Single Peaks

Parameter	Value
Max Allowed Eccentricity	0.7
Max Area	Mean of Area of Blobs in B_i

Table 5.3: Criteria for determining if a blob represents a single peak.

The resulting conditioned and normalized image (Figure 5.5f) has barcodes with pixel intensities such that the brightest pixel in an average barcode in any channel will have an intensity of one. Together with information about the locations of the barcodes (Section 5.4.1), this normalized and conditioned image can be used to create one-dimensional barcode intensity profiles (Section 5.4.1) that can then be decoded using the framework described in Section 5.3.3.

5.4.3 Inspecting the Background

The decoding framework described in Section 5.3.3 is designed to operate with zero-mean Gaussian background noise. We verify the nature of the noise by masking away the prospective barcodes using the inverse of the binary mask $B_{\cup}$ from Section 5.4.1. The intensity histogram of the background, shown in Figure 5.6, demonstrates that the noise is close to zero mean. The same data plotted on normal probability axes shows that over 90% of the noise pixels fit a Gaussian distribution. The remainder have lower intensities than expected, so treating them as Gaussian is conservative in the sense that they are unlikely to cause false positives by masquerading as fluorescent markers. The background noise is thus well described by a zero mean Gaussian distribution, and is therefore consistent with the assumptions in our model.

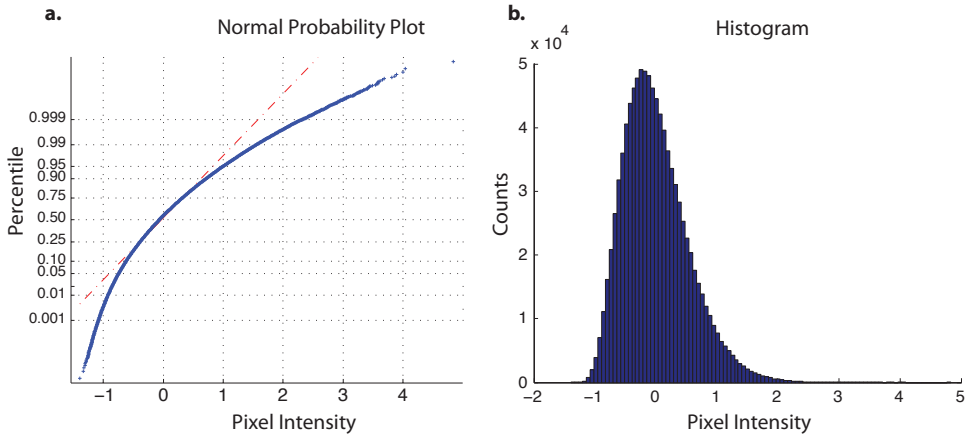


Figure 5.6: The background of the normalized and conditioned image is roughly zero mean and Gaussian. **(a)** A normal probability plot compares the cumulative distribution function of the pixel intensities (blue "+" marks) of the background of the red channel of a normalized condition image containing barcodes with that of a Gaussian (red dashed line.) **(b)** Histogram shows number of pixels for given intensity ranges. For both **a** and **b** the background pixels were collected from the normalized conditioned image by masking away the location of the barcodes and any other bright fluorescent features.

5.4.4 Projecting the Barcode to One Dimension

To decode a barcode using the one-dimensional matched filter described earlier, we need to first project each three-channel barcode image into three one-dimensional intensity profiles $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$. We use the location and orientation of the barcodes found in Section 5.4.1 to rotate the barcode, draw a rectangular bounding box around the barcode and to project it down to a single profile, see Figure 5.7.

For a given barcode, we first rotate each channel of the conditioned image $U_i(x, y)$ around the barcode's centroid until the major axis of the barcode lies along the x axis. We call this rotated image $U_{\theta,i}(x, y)$. The one-dimensional projection of a barcode in the ith color channel is then,

$$\mathbf{u}_i = u_i(x) = \frac{1}{2\Delta y \kappa \langle g_i \rangle} \int_{y_o - \Delta y}^{y_o + \Delta y} U_{\theta,i}(x, y) \mathrm{d}y, \tag{5.26}$$

where $\langle g \rangle_i$ is the average peak height for this channel described in Section 5.4.2; the rectangular region is centered vertically at y_o and is $2\Delta y$ tall; and κ is a correction factor, the same for all channels, so that the average peak height remains unity even after the projection. κ is the integral of the point spread function h in the vertical direction

$$\kappa = \int_{-\Delta y}^{\Delta y} h(y) \mathrm{d}y \tag{5.27}$$

where Δy is chosen to capture the majority of the energy within h. All rectangular regions have the same height Δy. Our analysis models the point-spread function h as a Gaussian with width σ_{psf},

$$h(r) = \exp\left(-\frac{r^2}{2\sigma_{\text{psf}}^2}\right), \tag{5.28}$$

so κ is

$$\kappa = \frac{\sqrt{2\pi}}{2\Delta y} \mathrm{erf}\left(\frac{\Delta y}{2\sqrt{2}\sigma_{\text{psf}}}\right) \tag{5.29}$$

where erf is the error function [131]

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_o^z \exp(-t^2)\, \mathrm{d}t. \tag{5.30}$$

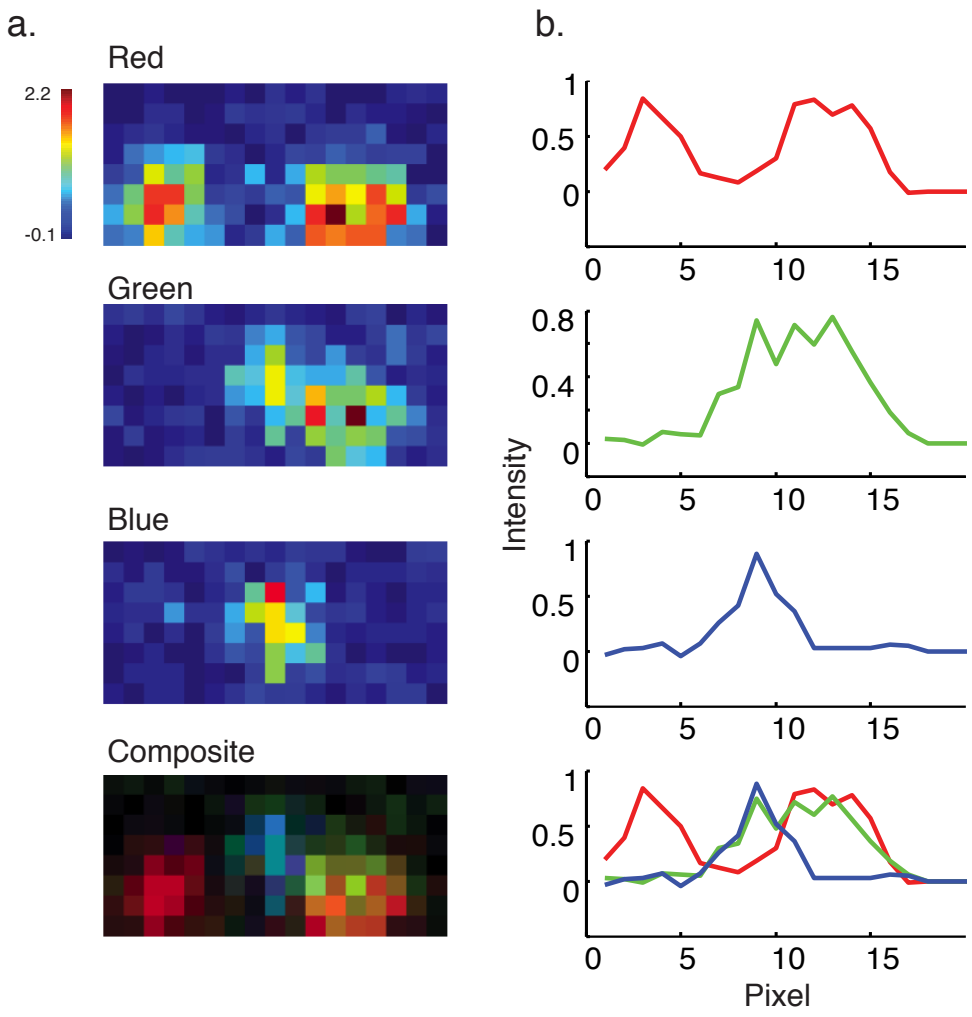


Figure 5.7: Two-dimensional barcode images are projected down to a single dimension. 1 px $\approx$ 71.4nm. **(a)** A single barcode is shown. From top to bottom, the Red, Green, Blue and a composite image are shown. These images have already been conditioned and normalized. **(b)** The one-dimensional projections of the images in **a** are shown.

Eq. 5.26 provides us with three one-dimensional projections of the barcode intensity with average peak heights of unity and with zero-mean Gaussian noise. These projections are now ready to be compared to a reference to decode the barcode.

5.4.5 Generating References

The last ingredient needed for detection and identification is a bank of reference barcodes to compare against the observed signal. Each reference barcode consists of three one-dimensional traces containing Gaussian peaks of height unity (Eq. 5.28) superposed at the predicted locations of the fluorescent spots.

Although there are only $6^3 = 216$ species of barcodes, the reference bank must be larger to accommodate the many differences in geometry, or microconfigurations, that the barcode can adopt, including:

1. left/right reflections, that is, the small gap can lie to the left or to the right,
2. left/right mis-centering within the rectangular window,
3. variation in the fluorophore spacings arising from the barcode manufacturing process, and
4. bends in the DNA strand resulting in a mild U rather than a linear shape.

These differences are accounted for by varying each barcode reference waveform in half-pixel increments, so that the observed profile $\mathbf{u}$ is compared to a bank of references encompassing each possible geometric microconfiguration. Excluding left/right mis-centerings, there are 48 variations that each of the 216 barcode species might adopt, for a total of 10,368 three-channel waveforms in the reference bank. Table 5.4 lists the parameters that are allowed to vary with their ranges.

Left/right translations are handled by a circular convolution operation described in the following section. Note that we are not accounting for differences in fluorescence a_j caused by inefficient labeling during manufacturing, because that would be too computationally intensive. For now, we instead pre-screen barcodes to ensure that they are properly labeled, as discussed in Section 5.7. As

Reference Barcode Properties

Parameter	Value
Point Spread Function Width	$\sigma^2 = 1.59$ px
Allowed Lengths for Big Gap	{5.5, 6, 6.5, 7, 7.5, 8} px
Allowed Lengths for Little Gap	{ 3.5, 4, 4.5, 5 } px
Window Width	64 px
Number of Barcode Species	216
Number of Microconfigurations	10,368

Table 5.4: Reference barcode properties. Scale: 1 px $\approx$ 71.4 nm.

DNA origami manufacturing techniques improve over time, this step will become unnecessary.

Once the bank of references is generated, it is then ready to be used in the multiple hypothesis test to determine the most likely barcode corresponding to a given observed signal, as described below.

5.4.6 Decoding the Barcode

The image processing steps described thus far have prepared the signals for their eventual decoding with a multiple hypothesis test, which is the main task of the software and which we now describe. We test a given observed signal $\mathbf{u}$ composed of three one-dimensional traces as in Eq. 5.26, against all 10,368 references, using Eq. 5.23 to identify the most likely one. Expanding this equation for the ith channel gives

$$\max_k \{L_k\} = \max_k \left\{ \sum_{i=1}^{3} \frac{2\mathbf{u}_i^T \mathbf{m}_{ki} - \|\mathbf{u}_i\|^2 - \|\mathbf{m}_{ki}\|^2}{\sigma_i^2} \right\}, \tag{5.31}$$

where $\mathbf{m}_{ki}$ is the kth reference and σ_i the noise, both in the ith channel. Note that $\|\mathbf{u}_i\|^2$ is independent of the reference k being used for comparison and can be omitted from the maximization, so that

$$\max_k \left\{ p(\mathbf{m}_k|\mathbf{u}) \right\} = \max_k \left\{ \sum_{i=1}^{3} \frac{2\mathbf{u}_i^T \mathbf{m}_{ki} - \|\mathbf{m}_{ki}\|^2}{\sigma_i^2} \right\}. \tag{5.32}$$

The term $\mathbf{u}_i^T \mathbf{m}_{ki}$ is a convolution that, by definition, is a matched filter for the kth reference waveform. This equation may be interpreted, therefore, as a multi-hypothesis bank of matched filters [72].

Our work would be complete at this point, were it not for possible left/right mis-centering of the barcodes within the rectangular window, as mentioned in Section 5.4.5. Instead of adding additional references to account for all possible translations of all 10,368 references so far, we use the property that translations in the barcode correspond to x offsets or *lags* in the convolution $\mathbf{u}_i^T \mathbf{m}_{ki}$ between the reference and signal. Symbolically we write this as

$$\max_k \{L_k\} = \max_{k,x} \left\{ \sum_{i=1}^{3} \frac{2(\mathbf{u}_i * \mathbf{m}_{ki}) - \|\mathbf{m}_{ki}\|^2}{\sigma_i^2} \right\}, \tag{5.33}$$

where $*$ denotes convolution. This is maximized for the best reference $\mathbf{m}_k$ and for the maximum value of convolution, which occurs for the lag that produces the largest result. It is computationally efficient to calculate such a convolution for all values of lag by employing the Fast Fourier Transform algorithm. The observed one-dimensional profile $\mathbf{u}_i$ is first zero-padded to equal the length of the reference. It and the reference $\mathbf{m}_k$ are then transformed, their spectra multiplied and the product inverse transformed to find the result [122]. Moreover, both $\|\mathbf{m}_{ki}\|^2/\sigma_i^2$ and the Fourier transform of each reference m_k can be calculated once in advance, and stored to reduce analysis time. These stored values may then be used to test all barcodes in the image.

5.5 Sources of Error

5.5.1 Barcode Defects

DNA origami barcodes are a nascent technology that will surely see rapid improvements in manufacturing robustness, quality and uniformity. The initial protocol for barcode manufacture as described in Chapter 4 produces numerous barcode defects, predominantly of two types. First, we see barcodes that have the wrong number of fluorescent spots, either only two spots or four spots or more. This is likely the result of a barcode that folded improperly or of two barcodes that became conjoined together. The software usually rejects these barcodes based on

their morphology, but occasionally the defect conspires to have roughly the same shape as a regular barcode and thus the algorithm is fooled into decoding them. Secondly, we see a more subtle defect. Frequently, two spots in the same color channel on one barcode will vary dramatically in intensity, by as much as four-fold. Careful inspection by eye reveals a very bright and very dim spot. This is likely caused by poor labeling of the active site. Each site or spot should have six fluorophores, but we suspect that often the number of fluorophores is less. We evaluate the effectiveness of our software both including and excluding these two classes of defects.

5.5.2 Imaging Conditions

Our model does not take into account stage drift that may cause slight misalignment between frames nor sparodic changes in laser intensity on a ~1 min time scale. Changes slower than ~1 min are corrected for by taking new regular background and blank images.

5.6 Software

The software consists of MATLAB scripts used with MATLAB Version 7.11.0.584 (R2010b) [114]. The software calls functions from the following toolboxes: Image Processing Toolbox Version 7.1, Fuzzy Logic Toolbox Version 2.2.12 and Signal Processing Toolbox Version 6.14. The analysis software scripts are released under the GNU General Public License and are available to download from `http://github.com/aleifer/DNAbarcode`.

5.7 Results

To test the system, we compared the software's identification of a pool of barcodes to that of a human. The human performed the identification and recorded the results before the software so that the human was not influenced by the software's identifications.

We examined three images containing 189 fluorescently-labeled objects that match the morphological criteria for a barcode shown in Table 5.2. Of those 189,

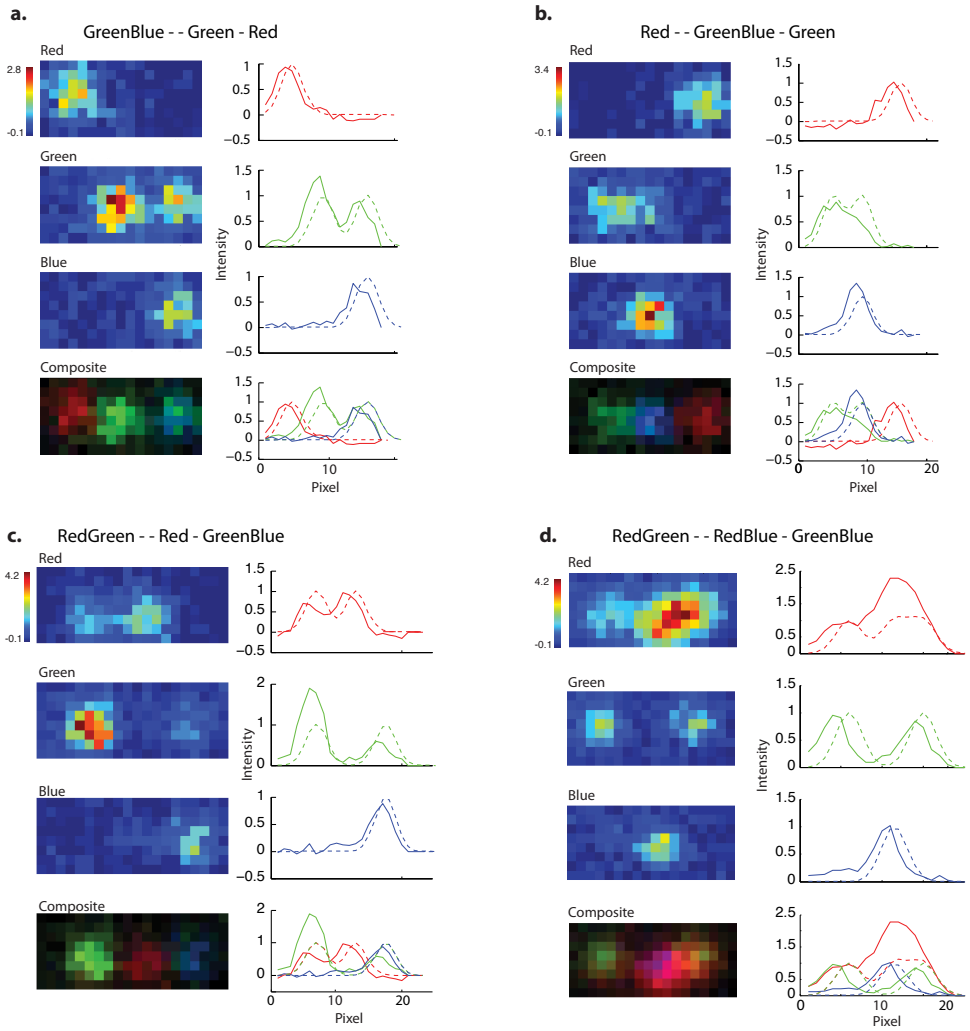


Figure 5.8: The software decodes barcodes (solid line) by correctly identifying the most likely reference (dashed line). Four representative barcodes are shown. In each case the software correctly decodes them. The barcodes in **(a)** and **(b)** are well formed and have consistent fluorescent labeling (peak heights fall between 0.75 and 1.5). The barcodes in **(c)** and **(d)** have peaks that vary widely, even within a channel. This is likely caused by poor labeling during the manufacturing process.

we manually selected only those barcodes that were fully formed and evenly labeled. To be fully formed the barcodes must contain three fluorescing spots. To be evenly labeled, all peaks in a given barcode must have a height that falls in a range between 0.7 and 1.5. (Recall that the images have been normalized so that the mean peak height in each channel has intensity of unity.) For example, the barcodes in Figure 5.8a and b have even fluorescent labeling and match the criteria, while the barcodes in Figure 5.8c and d have uneven labeling. Of the 189 putative barcodes in the three images, 28 (15%) fit the criteria for uniform labeling. The computer decoded these 28 barcodes and matched the human expert 27 out of 28 times, giving an accuracy of approximately 96%.

Although it was designed to operate on uniform, well-formed barcodes in a production setting, the software can still handle non-ideal cases. Even in extreme cases when the barcodes contain peak heights that are grater than 1.75 or less than 0.5, the software still succeeds 31% percent of the time (9 out of 29 barcodes, examined in a fourth image).

5.8 Discussion

We have presented mathematical theory and software implementation to locate DNA origami barcodes in an image and decode them.

We locate the barcodes based on morphology, reduce the two-dimensional images of the barcodes down to one-dimensional profiles and then use Bayesian multiple hypothesis testing to decode the barcodes. We show that a Bayesian multiple hypothesis test of a bank of matched filters is an optimal method of decoding the barcodes. We also model such real-world imaging conditions as inhomogeneous illumination, camera offset and background fluorescence.

The software works very well at decoding barcodes that have been pre-screened to have uniform fluorescent labeling. While only a small percentage (~15%) of manufactured barcodes currently meet our criteria for uniform labeling, we are confident that the labeling efficiency will increase dramatically in coming years. One promising approaches is to add crosslinkers to the barcodes after they are fluorescently labeled to prevent fluorescent DNA oligos from falling off.

5.9 Manuscript Information

5.9.1 Submitted for Publication As

A version of this chapter is being submitted for publication to the journal *Bioinformatics*.

5.9.2 The Author's Contribution

Andrew M. Leifer wrote all of the software, generated all figures, developed portions of the mathematical framework and wrote the majority of the manuscript. Mark C. Leifer developed portions of the mathematical framework and wrote portions of the manuscript. Chenxiang Lin created the DNA origami barcodes and conducted all microscopy work.

6

Neural activity of the Omega Turn

Recently, the change in the neural activity of the *C. elegans* motor circuit that occurs between the worm's forward and reverse locomotion has been an area of intense study in laboratories around the world [9, 47, 83, 126]. Three groups published work on this area in only the past few months [47, 83, 126]. In the ongoing work presented here, we study the neural activity that underlies the omega turn, which also includes two such transitions between forward and reverse locomotion. We confirm some of the results found by others and use our rich and detailed behavioral readout to probe more deeply at correlations between neural activity and specific subtle behaviors critical for navigation.

6.1 Motivation

Navigation is a goal directed locomotion, common across species, in which an organism moves toward or away from a cue. At the neuronal level, navigation requires the temporal coordination of different motor programs. How does an animal's nervous system orchestrate the changes in locomotion required for naviga-

tion? The *Caenorhabditis elegans* escape response provides a platform upon which to investigate how navigation is carried out by a simple nervous system.

Navigation in *C. elegans* is regulated by ventrally or dorsally biased head swings [76] and deep ventral omega turns that reorient the worm in the opposite direction. When the worm is touched on its anterior, it exhibits an escape response by pausing, reversing and turning ventrally in what is often called an omega turn. The neural circuit for this escape response employs cells at all levels of the nervous system: mechanosensory neurons, command neurons, motor neurons, and muscle. The circuit plays an important role in the animal's decision making, the coordination of its motor programs and its turning behavior. The analysis of this stereotyped omega turn provides an opportunity to identify the neuronal mechanisms that the nervous system employs to translate sensory information into navigational behavior.

In this chapter we present ongoing work to explore the neural activity underlying the omega turn. A real-time tracking system was developed to record intracellular calcium transients in single neurons while simultaneously monitoring the macroscopic behavior of a freely moving worm as it undergoes an escape response.

6.2 Dual-Magnification Calcium and Behavior Imaging

6.2.1 Background

To study the neural activity of an omega turn requires a system that permits simultaneous observation of neural activity and behavior in freely moving worms. In particular, it is important to be able to correlate individual neural activity to rich details of the worm's behavior, such as changes in curvature or direction.

Genetically encoded fluorescent calcium indicators such as Cameleon [115] and GCaMP [153] are the primary tools available for optical neurophysiology in *C. elegans*. The fluorescence of these indicators change with the level of intracellular calcium present. Calcium is often a good indicator of neural activity. As a result, the fluorescent levels can be monitored and neural activity can be inferred.

Calcium imaging is often performed at high magnification (20x or more) to resolve individual neurons and with high numerical aperture (NA) objectives to collect large numbers of photons. As a result, calcium imaging had traditionally been

performed on immobilized [29] or partially restrained animals [46], where it is possible to keep the neurons of interest in a small field of view for an extended period of time.

The Samuel Lab pioneered the first system to perform calcium imaging in a moving worm [30]. In that system, the user manually adjusted a joystick to track the neuron AFD in the head of an unrestrained worm on a motorized stage. The worm was imaged with a high NA 20x objective. This system allowed the user to observe calcium transients in a worm as it moved. Because the field of view only included a small portion of the worm, details of the worm's behavior like its curvature were not visible. Additionally, manual tracking made it difficult to consistently track abrupt changes in the worm's motion.

Since then, a number of groups have developed manual [83] and automated [9, 47, 126] tracking systems for calcium imaging in moving worms, using either home built [9, 83, 126] or commercial [47] systems. These systems vary in the feedback rate, complexity, cost and in their ability to capture behavior data. For example, the systems presented in [126] and [83] lack the ability to view the entire worm in the field of view.

We have a developed a new tracking system that offers unique advantages. The system is built around a spinning disk confocal microscope which cuts down on background fluorescence and increases signal to noise. The system identifies target neurons from the worm's outline, rather than from their fluorescence which makes it well suited for transgenic animals that exhibit fluorescence in more then one neuron. Most significantly, the system builds on previous experience developing rich behavioral monitors and thus provides high spatiotemporal behavioral data at 30 Hz including the worm's instantaneous curvature and orientation.

6.2.2 The DualMag System

A tracking microscope was built capable of recording calcium transients in a moving worm. The microscope operates at two magnification levels simultaneously, one beam path operates at high-magnification to resolve single neurons, while the other beam path operates at low magnification to view the entire body of the worm. Real-time computer vision software monitors the position of the body of the worm from the low-magnification beam path and adjusts a motorized stage to

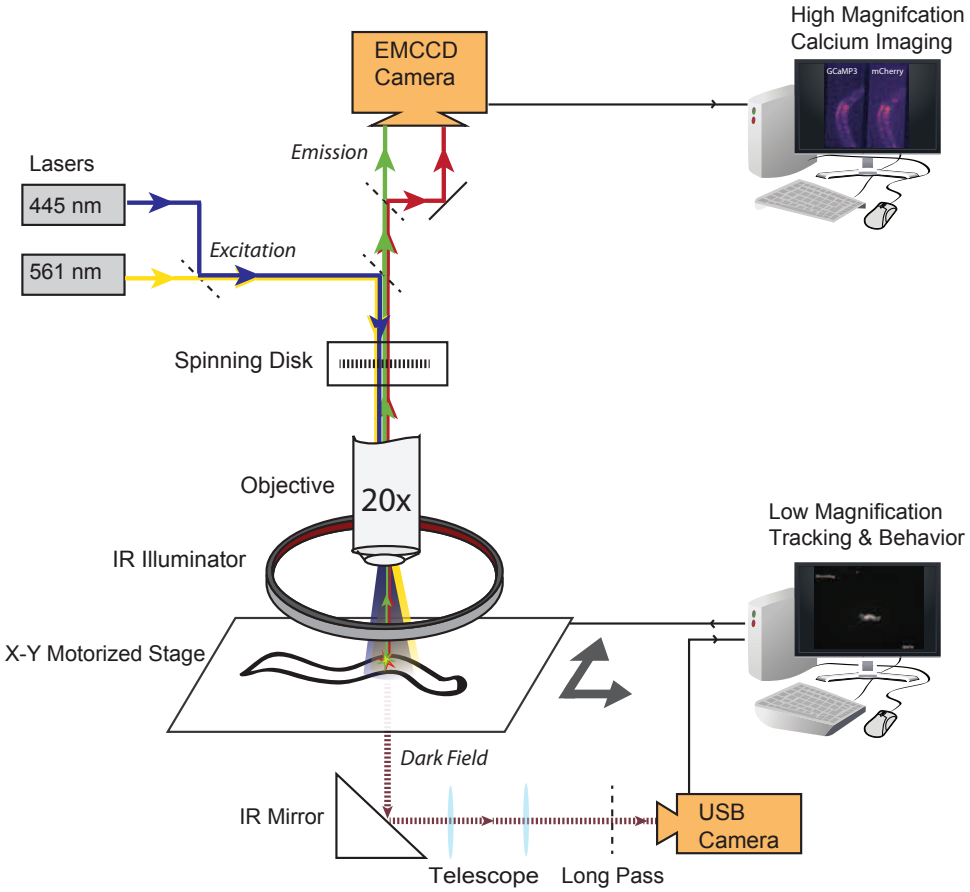


Figure 6.1: Schematic of the DualMag setup, permitting simultaneous recording of intracellular calcium transients and behavior in freely moving *C. elegans*. A transgenic worm, expressing GCaMP3 and mCherry in targeted neurons, crawls freely on a motorized stage under infrared illumination. An inexpensive USB camera images the worm's motion at low magnification. Real-time computer vision software rapidly analyzes each low magnification image and identifies the location and orientation of the worm and the targeted neuron and adjust the stage to keep the targeted neuron centered beneath a 20x objective used for calcium imaging. Blue and yellow laser light shine through the 20x objective to excite GCaMp3 and mCherry in the targeted neuron. The emitted green and red fluorescence is imaged through a spinning-disk confocal microscope onto two halves of an electron multiplying CCD camera. Comparing the fluorescence in the green and red channel images gives the relative level of calcium in the neuron. Slanted dashed lines indicate dichroic mirrors.

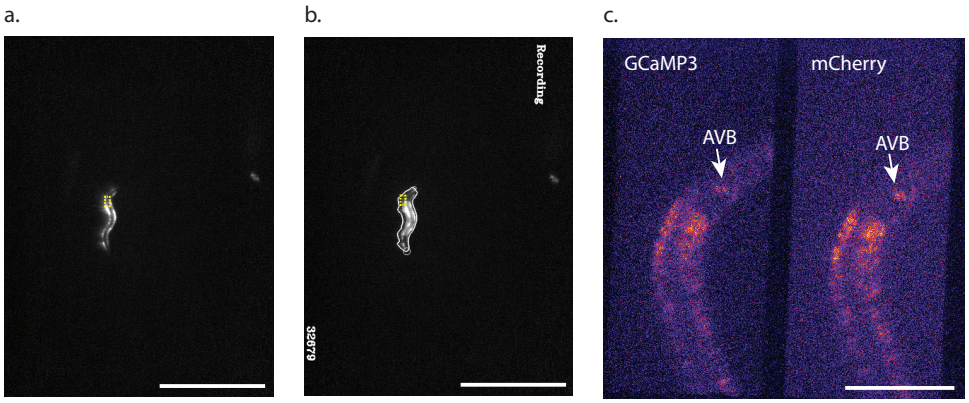


Figure 6.2: Example images of the DualMag setup. Transgenic worm expressing GCaMP3 and mCherry in the AVB interneuron is observed in the DualMag system. **(a)** Raw-image of worm behavior recorded by the low-magnification beam path is shown. Scale bar is 1mm. **(b)** Real-time computer vision software identifies the outline of the worm, its head and tail, and identifies the position of AVB for tracking. Scale bar is 1mm. **(c)** The high magnification beam path images the yellow square region shown in **b**. Green channel (left) shows calcium-dependent GCaMP3 fluorescence. Red channel (right) shows calcium-independent mCherry fluorescence. Scale bar is 100 µm.

keep targeted neurons in the high-magnification field of view. The entire system is called the DualMag.

The DualMag system is built around an upright microscope body with a spinning disk confocal unit, a motorized stage and an additional low magnification imaging path. See Figure 6.1 and Section 6.7.2 for details.

A transgenic animal expressing GCaMP3 and mCherry in targeted neurons crawls freely on agar in a petri dish on a motorized stage under dark field infrared illumination. The high magnification beam path illuminates targeted neurons with blue (445 nm) and yellow (561 nm) laser light though the spinning-disk confocal system, emitting fluorescence from GCaMP3 in the green and from mCherry in the red. Dichroic mirrors filter out the excitation light and pass only the emitted light to a DualView unit which projects the red and green channels side-by-side onto the sensor of an EMCCD camera. See Figure 6.2c and Section 6.7.2.

A second beam path, beneath the microscope images the infrared light scat-

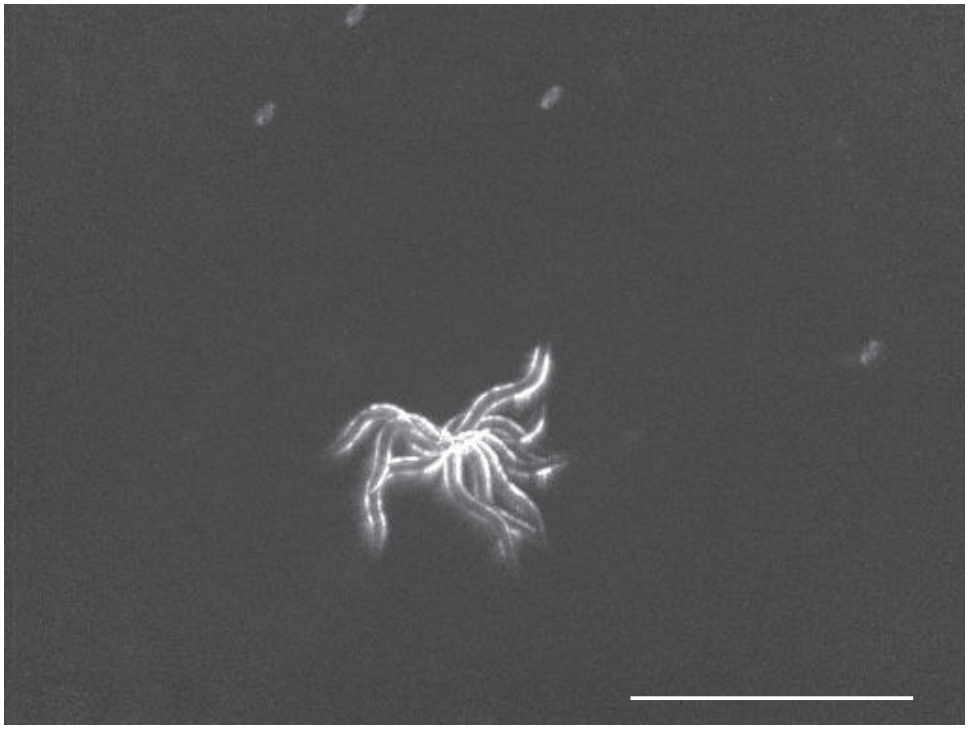


Figure 6.3: Time-lapse stroboscopic darkfield image of worm behavior from the low magnification beam path is shown. The DualMag system continually adjusts the stage's velocity to isolate motion in a region of the worm's head. A 28 ms exposure image was taken every 3 seconds for 1 minute. Scale bar is 1 mm.

tered by the worm at low magnification. An inexpensive USB CCD camera monitors the worm's position and orientation at 30 fps. See Figure 6.2a. Custom real-time computer vision software written in C identifies the worm's outline, its head and tail, and the location of targeted neurons (Figure 6.2b and Section 6.7.2) and instructs the motorized stage to adjust its stage velocity to compensate for the worm's motion and to keep the targeted neuron squarely in the center of the high-magnification field of view. The feedback loop operates at 30 Hz and is sufficient to compensate for the worm's natural movements. See Figures 6.3 and 6.4.

While similar in principal to prior work [9, 126] this is the first system, to our knowledge, that uses a spinning disk confocal microscope which cuts down on background fluorescence. Second, it uses an inexpensive USB camera to do the tracking and leverages the real-time computer vision software from the CoLBeRT

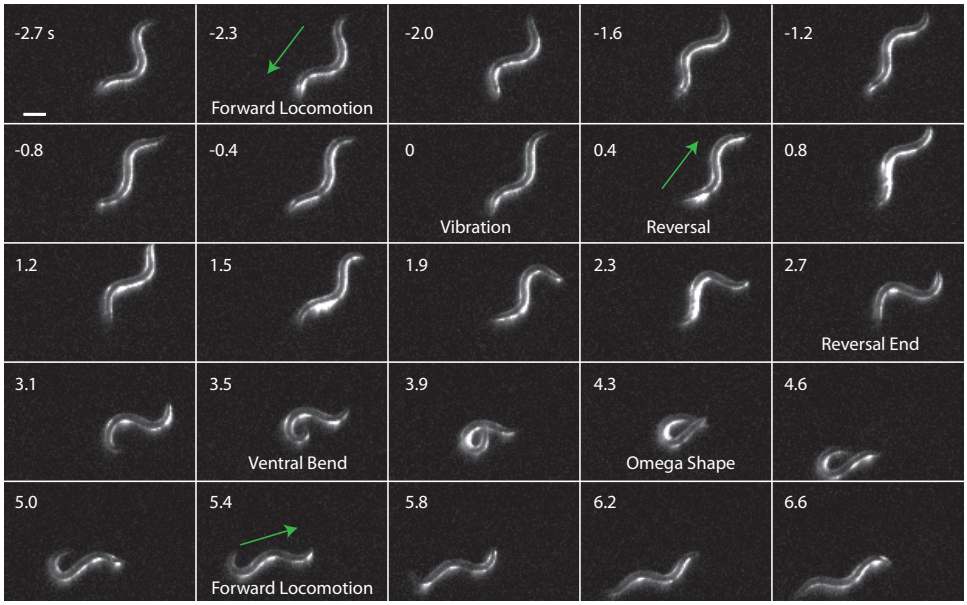


Figure 6.4: Video of worm escape response is shown. The worm undergoes forward locomotion. At 0 s, an electric toothbrush vibrates the worm's petri dish and the worm halts forward locomotion and begins a reversal. At 2.7 s the worm ceases reversing and begins a deep ventral bend. At 4.3 s the worm exhibits the omega shape that gives the omega turn its name. At 5.4 s the worm recommences forward locomotion. Scale bar is 100 µm.

system [92] to perform rapid feedback based on the worm's morphology, not individual neuron fluorescence.

6.3 Neural Activity of the Escape Response

The escape response in *C. elegans* can be elicited by gently touching the anterior of the worm or by vibration. The response is stereotyped: The worm first ceases its forward locomotion and exploratory head movements [3], it then moves backward away from the stimulus [22], comes to a stop, and then bends its head ventrally to execute a deep ventral turn before finally reinitiating forward locomotion. See Figure 6.4. The entire sequence is commonly referred to as an omega turn.

The escape response is initiated by mechanosensory neurons which transduce sensory information through locomotion command neurons to excitatory and in-

hibitory motor neurons that innervate the body wall muscles [22]. Laser ablation and genetic studies have elucidated which neurons are required for which behavior [3, 22, 176]. A wiring diagram representing current thinking in the field is summarized in Figure 6.5. Despite a rich literature describing which neurons are required for which behavior, it is only very recently that researchers have begun to study how neural activity correlates to behavior, and even then only in regards to forward versus backward locomotion [9, 47, 83, 126]. To fully understand the nervous system, one must also investigate the dynamics of neural activity and investigate the precise patterns and sequences of neural activity that drive behavior. As a first part of this ongoing work, the pattern of neural activity in the neurons AVA and AVB are investigated and correlated to the precise pattern of behavior exhibited by the worm during an omega turn.

6.4 Results

6.4.1 AVA

AVA is a command interneuron involved in backward locomotion. A P*rig-3::mCherry::SL2::GCaMP3* transgenic worm, expressing mCherry and GCaMP3 in AVA is stimulated to undergo an escape response by transiently vibrating an electric toothbrush against the side of the agar plate containing the worm for approximately a second. As the worm freely crawls on the agar, its behavior and calcium transients are recorded. A representative sequence is shown in Figure 6.6. After vibration, calcium levels in AVA begin to rise as the worm reverses. The worm ceases backward locomotion, and undergoes a deep bend, visible as a large downward spike in the curvature diagram. Its body then forms a symmetric omega-like shape and the worm begins forward locomotion in a new direction, evident by the jump in orientation.

The fluorescence ratio is the ratio of the of intensities, above background, of the pixels of AVA in the green to the red channel. Where presented, fluorescence ratio is plotted as the ratio above a baseline. The baseline is defined as the mean fluorescence ratio during a period from the beginning of a behavior sequence to

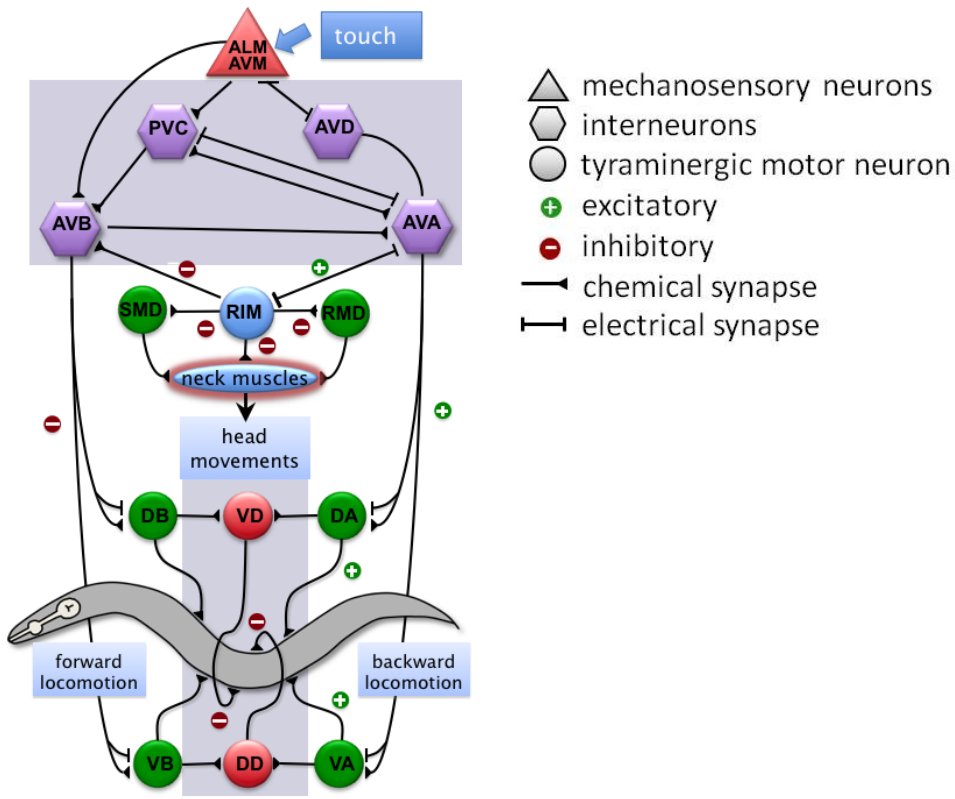
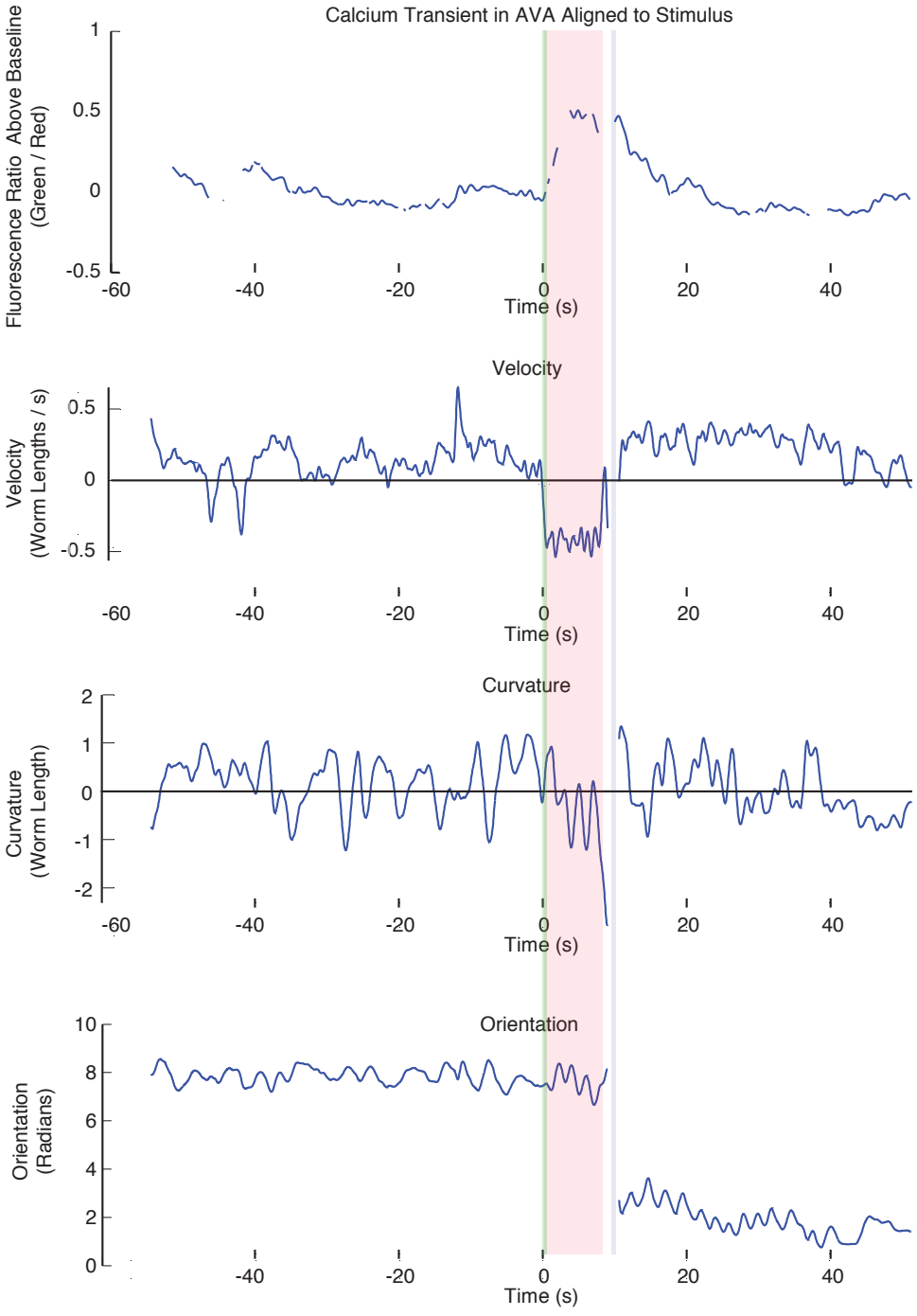


Figure 6.5: Circuit diagram showing neurons and their suspected role in the escape response. Information shown here is based on cumulative evidence from laser ablation and genetic studies performed by the community. Adapted from [3] and [127].

Figure 6.6 ***(following page)*:** Calcium transient of AVA is shown with the worm's behavior. A P*rig-3::mCherry::SL2::GCaMP3* transgenic worm, expressing mCherry and GCaMP3 in its AVA interneuron is stimulated at time 0 s by vibrating its agar plate (light green background). The worm ceases locomotion and reverses (light red background) and then forms an omega turn (light blue backround) before reinitiating forward locomotion. Changes in velocity, curvature and orientation corresponding to an omega turn are clearly evident. Calcium levels in AVA increase as soon as the worm begins reversing and begin falling after re-initiation of forward locomotion.

Figure 6.6: (continued)



the application of stimulus. Here, a behavior sequence is a defined as a region of behavior exhibiting only forward locomotion punctuated by exactly one omega turn in response to a stimuli.

Where presented, velocity is the velocity of the worm's bending waves along its body, in units of worm length. Curvature is the curvature of the worm's centerline in a region 10 to 80 percent along its centerline from the anterior of the worm. A gap in the fluorescence trace indicates either an instance where the worm moves its head out of plane and the microscope transiently loses focus, or an instance when the tracking software incorrectly segments the worm and the neuron leaves the field of view. A gap in behavior data reflects an instance when the software is unable to correctly identify the worm's head and tail, as occurs when the worm curls up in the omega shape.

Neural activity and behavior have modest variation across trials and worms, but is overall stereotyped. The calcium transients in AVA of seven runs from four worms are shown in Figure 6.7. The worm's velocity, curvature and orientation corresponding to these traces are shown in Figures 6.8, 6.9 and 6.10, respectively.

6.4.2 AVB

AVB is a command interneuron suspected to be involved in forward locomotion. Analagously to with AVA above, a P*lgc-55::mCherry::SL2::GCaMP3* transgenic worm expressing GCaMP3 and mCherry in the neuron AVB is stimulated by brief vibration. The worm undergoes an escape response and its calcium transients and behavior is measured. See Figure 6.11 for a representative trace.

Prior reports [83, 176] have suggested that AVB is active during forward locomotion. In our hands, AVB appears to become active after the worm completes the omega turn and begins forward locomotion. Preliminary results, however, suggest that AVB activity may turn off some length of time (~30 seconds) after stimulus. Certainly in the trace shown here, AVB is inactive significantly before and after stimuli when the worm is undergoing forward locomotion. Perhaps AVB

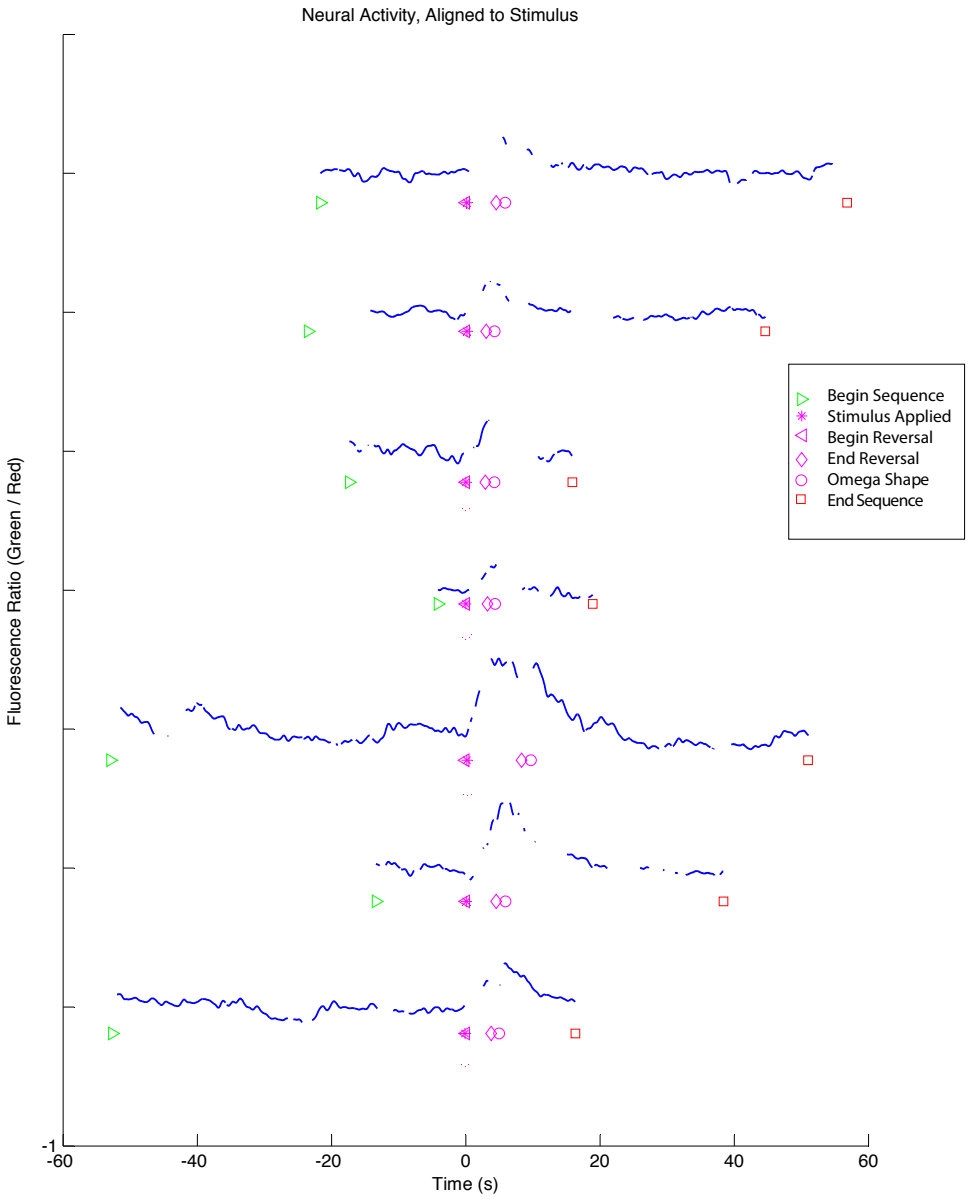


Figure 6.7: Calcium transients of AVA for seven behavior sequences from four worms. A P*rig-3::mCherry::SL2::GCaMP3* transgenic worm, expressing mCherry and GCaMP3 in its AVA interneuron, is stimulated at time 0 s by vibrating its agar plate. The fluorescence ratio is the ratio of the of intensities, above background, of the pixels of AVA in the green to the red channel, plotted here as the ratio above baseline. Corresponding velocity, curvature and orientation are shown in Figures 6.8, 6.9 and 6.10, respectively.

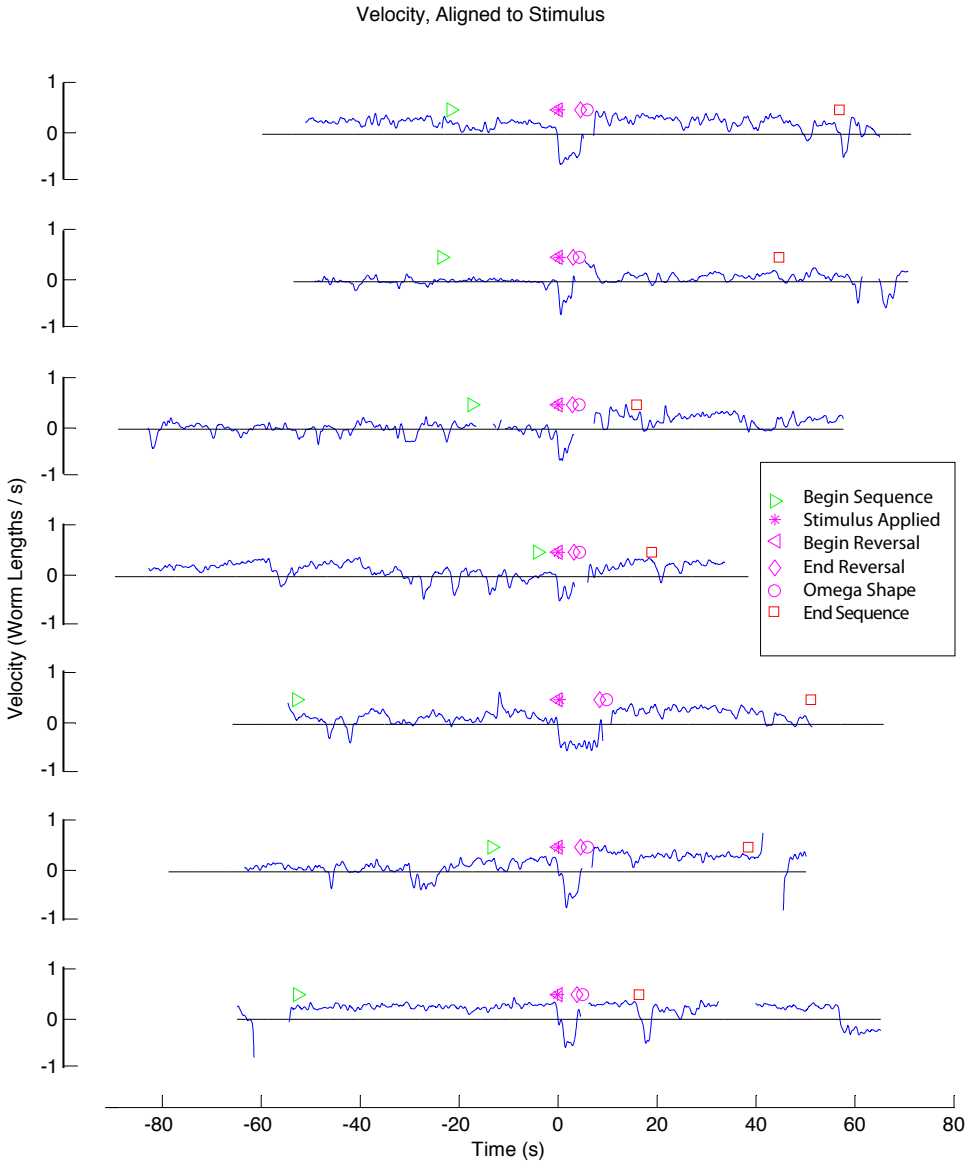


Figure 6.8: Velocity traces corresponding to AVA calcium traces shown in Figure 6.7. A P*rig-3::mCherry::SL2::GCaMP3* transgenic worm, is stimulated at time 0 s by vibrating its agar plate. Sustained negative dips correspond to reversals. Velocity, here, is the velocity at which the bending waves propagate along the worm's centerline. The worm's corresponding curvature and orientation are shown in Figures 6.9 and 6.10, respectively.

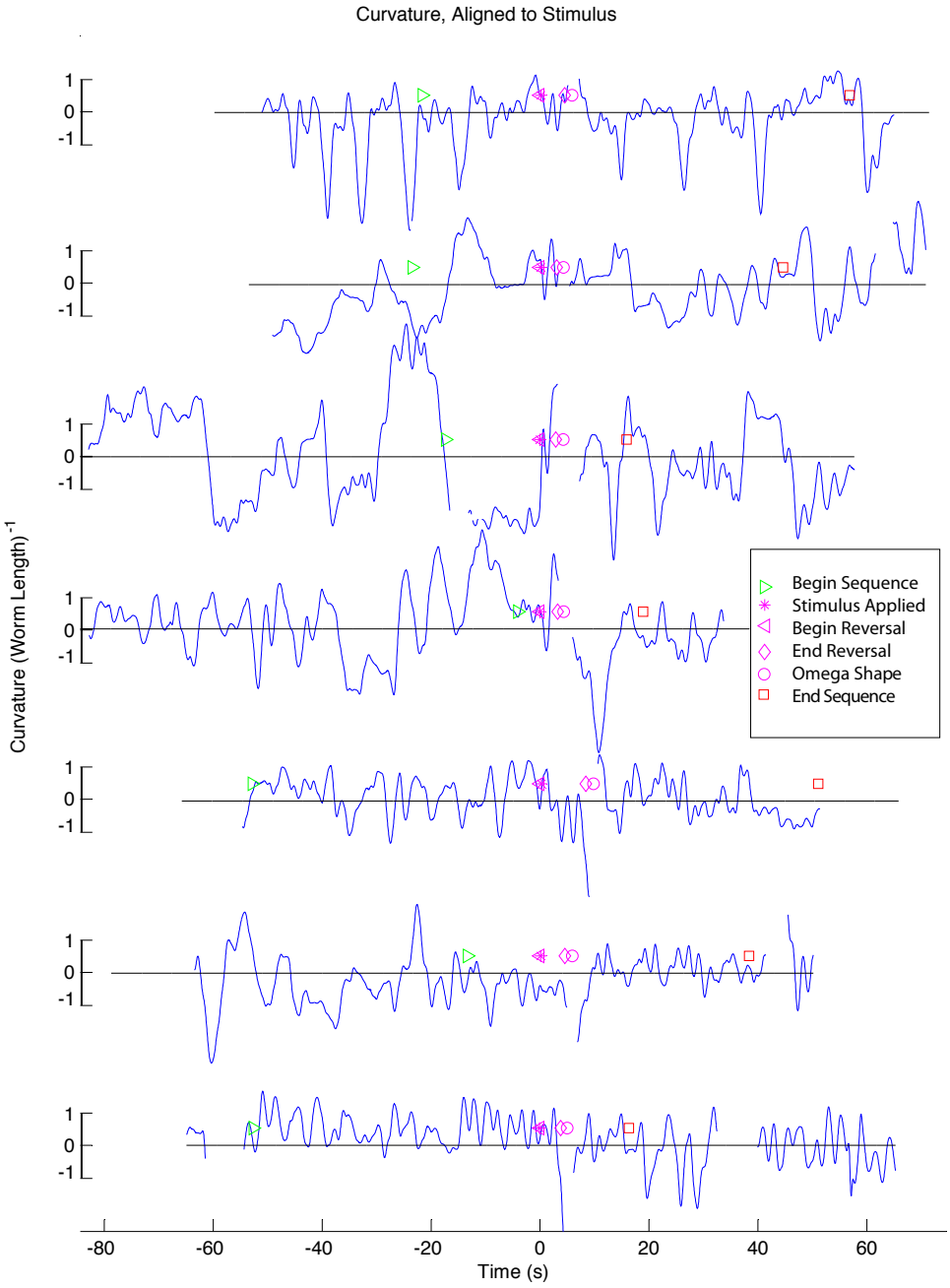


Figure 6.9: The worm's curvature ,corresponding to AVA calcium traces shown in Figure 6.7. A P*rig-3::mCherry::SL2::GCaMP3* transgenic worm is stimulated at time 0 s by vibrating its agar plate. A large spike or dip following the end of the worm's reversal (pink diamond) corresponds to a deep ventral bend. Here, curvature is the mean curvature of the worm's centerline in a region from 10% to 80% along its centerline. Velocity and orientation corresponding to these traces are shown in Figures 6.8 and 6.10, respectively.

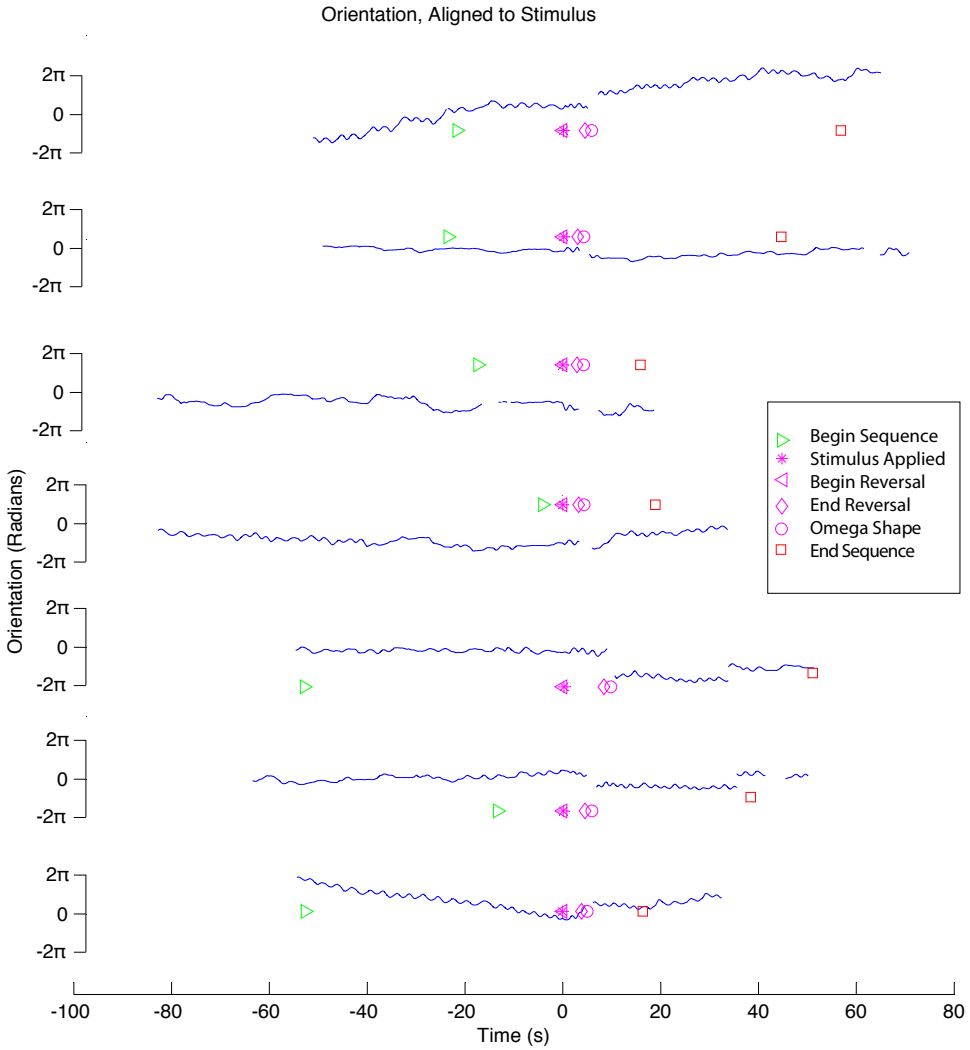
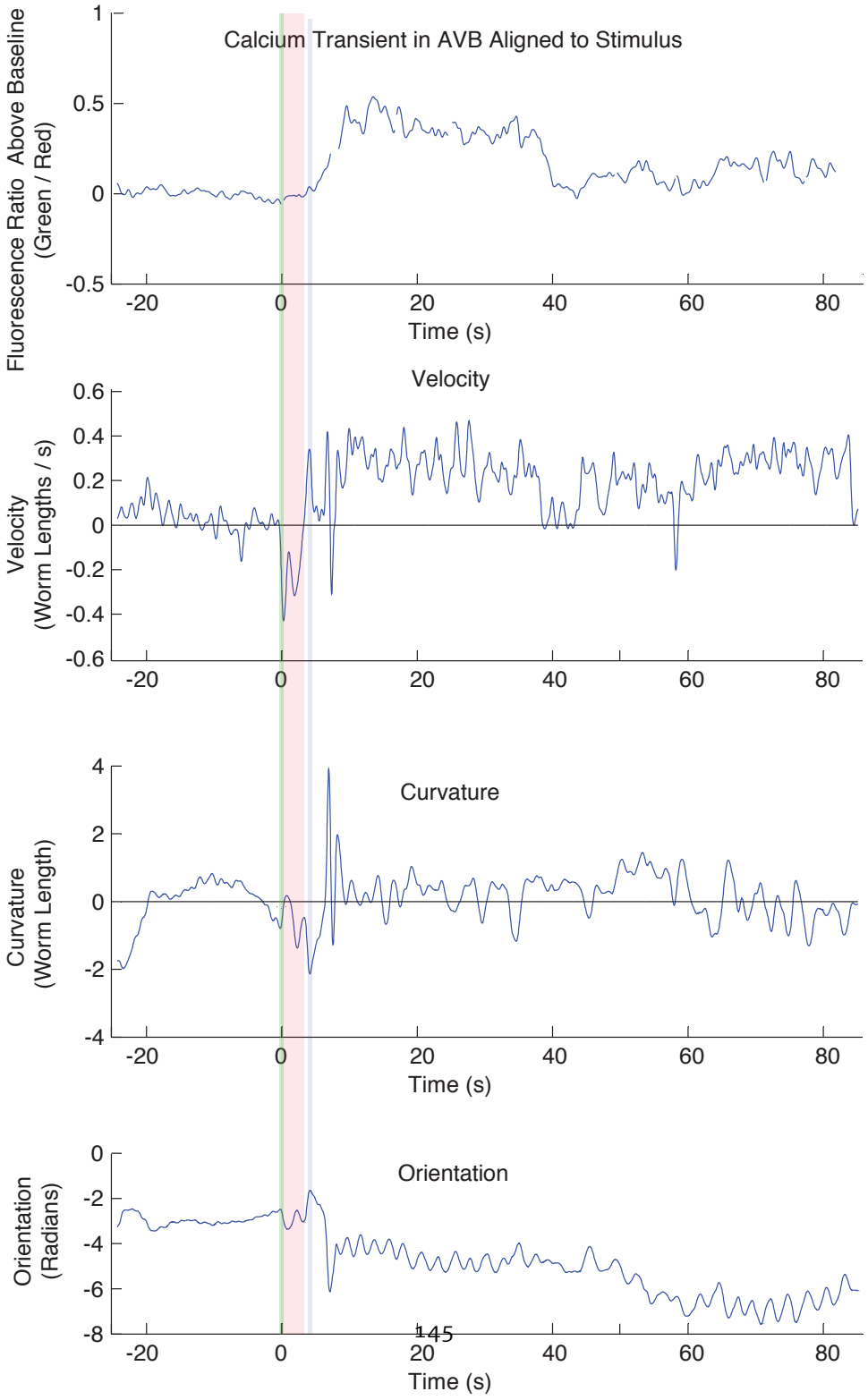


Figure 6.10: The worm's orientation, corresponding to AVA calcium traces shown in Figure 6.7. A P*lgc-55::mCherry::SL2::GCaMP3* transgenic worm is stimulated at time 0 s by vibrating its agar plate. A discontinuous jump around the time of omega-turn (pink circle) corresponds to an abrupt change in direction of motion. Here, orientation is the vector between the worm's neck and mid-region (5% and 40% along the from worm's centerlin anterior from its head). Velocity and curvature corresponding to these traces are shown in Figures 6.8 and 6.9, respectively.

Figure 6.11 ***(following page)*****:** Calcium transients indicating neural activity of AVB is shown with the worm's behavior. A P*lgc-55::mCherry::SL2::GCaMP3* transgenic worm, expressing mCherry and GCaMP3 in its AVA interneuron is stimulated at time 0 s by vibrating its agar plate (light green background). The worm ceases locomotion and reverses (light red background) and then forms an omega turn (light blue backround) before reinitiating forward locomotion. Changes in velocity, curvature and orientation corresponding to an omega turn are clearly evident. Calcium levels in AVA increase as soon as the worm begins reversing and begin falling after re-initiation of forward locomotion.

Figure 6.11: (continued)



is involved in maintaining a sprint or some other sort of heightened state of activity immediately following reorientation. More traces need to be recorded before any conclusions can be drawn.

6.5 Discussion

Over a decade ago researchers used laser ablation studies to suggest the role that AVA and AVB might play in the motor circuit [176]. This work confirms the results seen by others [9, 47, 83, 126] that AVA is active during reversals. This work also observes AVB activity. Preliminary results suggest that while AVB activity may be active with forward locomotion in response to a stimuli, it may not be persistently active during forward locomotion at all times. However, more recordings need to be performed to draw any conclusions.

This ongoing work also provides a rich quantitative behavioral readout of the behavioral sequence of the escape response and omega turn. Features such as reversals, pauses, bends and reorientations can be precisely quantified by measuring the worm's velocity, curvature and orientation. When these metrics are recorded simultaneously with calcium transients, it then becomes possible to make correlations between neural activity and behavior.

6.6 Future

This work is ongoing. The instrumentation required to simultaneously monitor calcium activity and behavior is complete and has already been used to probe the command interneurons AVA and AVB, although more trials are needed. In addition to recording more traces from AVB, the authors are also in the process of observing the interneuron RIM, where there is currently controversy about the role it plays during a reversal.

Eventually, this project aims to record calcium transients from all of the different neurons suspected to play a role in the escape response (Figure 6.5) and explore their activity in relation to the worm's behavior. The goal is to develop a picture of which neurons turn on and off when in the sequence of the escape response. Then a perturbative investigation can be performed, using tools like the CoLBeRT system (Chapter 2) to interrogate the individual contributions that silencing or

stimulating each neuron gives to behavior. This will be a long-term project, but the groundwork has been laid here.

6.7 Methods

6.7.1 Strains

The *C. elegans* strains used include QW280: *zfIs12[Prig-3::ChR2::GFP, lin-15(+)]* , QW625: *zfIs42[Prig-3::GCaMP3::SL2::mCherry; lin-15(+)]*, QW663: *lin-15(n765ts); zfEx219[Pcex-1::GCaMP3::SL2::mCherry; lin-15(+)]*, QW666: *lin-15(n765ts); zfEx222[Plgc-55::mCherry::SL2::GCaMP3; lin-15(+)]*. Transgenic strains were generated by microinjection of plasmid DNA along with the *lin-15* rescuing construct pL15EK into the germline of *lin-15(n765ts)* mutants. ChR2, GCaMP3 and *lin- 15(+)* plasmids were injected at 100 ng/µl, 50 ng/µl and 80 ng/µl respectively. Extrachromosomal arrays were integrated into the genome by X-ray irradiation (120kV) for 10 minutes. Integrated strains were outcrossed four times to N2 wild type.

6.7.2 Calcium Imaging in Moving Worm

Optics

A spinning disk confocal microscope built on a Nikon Eclpse LV 100 upright microscope frame form a foundation for the DualMag calcium imaging system.

The spinning disk system is a CSU22 Confocal Scanning unit from Yokogawa that operates at 5000 rpm. Blue (445 nm) and yellow (561 nm) lasers are fiber-coupled into the spinning disk system. Lasers are housed in an Andor Laser Combiner System 5000, and are controlled by an Andor Precision Control Unit 100 from Andor Technolgy. Confocal images are passed through a DV2 Dual-View unit (MAG Biosystems) so that the red and green channel images are projected onto an an Andor iXon+ EMCCD camera under the control of Andor iQ 1.1.3 imaging software. Confocal images were captured at 20 fps with an exposure time of 50 ms.

A 20x objective was used for an effective field of view, prior to the Dual-View, of about 300 µm wide. At this magnification each laser provided a maximum power of 20mW/mm^2 at the sample. In the experiments described here, the blue laser

was run at 3 mW/mm^2 and the yellow laser was run at 10 mW/mm^2 to avoid photobleaching.

To image behavior, a custom-made illumination ring of 100 850nm LED's was used to create dark field illumination by shining light perpendicular to the main optical path.

The condenser lens was removed from the bottom of the microscope to make room for a second low-magnification beam path to record the dark field imaging. An infrared prism mirror reflects infrared light into an optical train composed of a telescope and a 750nm long pass filter (Thorlabs, FEL0750). An inexpensive USB CCD camera, DMK 31BU03 from The Imaging Source, was used to monitor the worm's behavior.

The microscope stage was controlled by a Ludl Bioprecision2 XY motorized stage and MAC 6000 stage-controller. The worm was kept centered by a feedback loop implemented by the real-time tracking software (see below).

Real-Time Tracking

Real-time tracking was performed by monitoring the image of the worm through the low-magnification dark field beam path and adjusting the velocity of a motorized stage. 1024 x 768 px grayscale images were acquired from a USB camera 30 fps. A modified version of the MindControl software (Chapter 2) was used to track the worm and keep targeted neurons centered in the field of view of the high magnification beam path. In brief, the software identifies the outline of the worm, finds the head tail, segments the worm into coordinates, then identifies the targeted neuron within the coordinate system. Finally, the software measures the distance of the targeted neuron from the center of the high magnification field of view and adjusts the stage velocity to glide the worm back towards the center.

The software is essentially the same as in Chapter 2 with the following minor modifications, which were necessary to address poor lighting and the demands for more rapid tracking: When identifying the outline of the worm the software now performs a dilation on the image [16] which can compensate for "holes" that appear in the worm due to lighting limitations. The newer version also allows the the outline of the worm to be smoothed directly, whereas the old version only allowed the underlying image be smoothed. Finally the new version of the software

provides finer control of the stage response to the feedback loop.

As before, the MindControl software was written in C using the OpenCV open source computer vision library [15, 16] with the hardware optimized Intel Integrated Performance Primitives.

Video Synchronization

The DualMag system records two independent video streams recorded at different frame rates on different computers. The two video streams are synchronized by shining synchronous pulses of LED light at the appropriate wavelength into both cameras. LEDs are controlled independently from the other software mentioned by a custom LabView program interfaced with a LabJack U3-HV digital to analog converter. The flashes of light are identified manually during post processing.

Ratiometric Fluorescence Analysis

Radiometric fluorescence analysis is performed offline in a quasi-manual process using custom scripts in MATLAB [114]. The user selects control points to align the red and green channel images to each other. Scripts then prompt the user to manually click on the neuron of interest for every second frame of fluorescent imaging. The software searches in a local region nearby each mouse click to find the brightest pixel. A circular binary mask containing 50 pixels is drawn centered around the mouse click. This defines the region of interest containing the neuron. The software interpolates to find the location of the neurons for the half of the frames that were not specified manually. For each frame, the software measure the mean intensity of the 20 brightest pixels in the region of interest in the red and green channels and subtracts off local background fluorescence. Background fluorescence is measured manually by finding the mean intensity of non-fluorescing regions of the worm in ImageJ [130].

Fluorescent ratios are calculated as the ratio of the pixel intensities of the red and green regions of interest in each frame above background. When plotted, a baseline is subtracted from the fluorescent ratio.

Fluorescent images are recorded at 20 fps. Time-series traces of the fluorescent ratios were smoothed by linearly interpolating over any gaps and convolution with a Gaussian of σ=5 frames (~250 ms) with padding of the end-point values. Finally

points that had previously been interpolated were removed.

Behavioral Analysis

Behavior analysis was performed using a modified version of the MindControl Analysis scripts (Chapter 2), written in MATLAB [114]. Velocity is defined as in Chapter 2. Curvature here, is defined as the mean curvature of the worm's body as defined in Chapter 2, but averaged over the region of the worm's centerline from 10% to 80% down from the anterior. Orientation is the arctangent of the vector defined by the points along the worm centerline 5% and 40% along the centerline from the anterior.

Time-series behavioral data was smoothed by linearly interpolating over any gaps and convolution with a Gaussian of σ=5 frames (~170 ms) with padding of the end-point values. Finally points that had previously been interpolated were removed.

6.8 Manuscript Information

6.8.1 Author Contributions

The work presented here was done by Andrew M. Leifer with Christopher Clark and Mark Alkema of the Alkema Lab. Andrew Leifer built the hardware with Mason Klein. Christopher Clark and Andrew Leifer together performed all of the experiments. Christopher generated all worm strains. Andrew wrote all software and analyzed the data. Laura Freeman built the video synchronization LED system.

Both Andrew and Christopher wrote the manuscript. Mark Alkema contributed Figure 6.5.

7

Conclusion

The development of new tools and techniques can drive scientific discovery when carefully applied. In this thesis, I developed a new instrument to non-invasively manipulate neural activity with single-neuron specificity in an intact freely moving worm. At the time of its creation, this was the first instrument of its kind. I also developed a second instrument to record calcium transients from a moving worm while simultaneously recording the worm's behavior. I used these tools to gain a better understanding of how the nervous system of the nematode *C. elegans* propagates bending waves and how it transitions from forward to reverse locomotion.

Both of the tools presented here rely on advances in the fields of optogenetics and computer vision. In particular, the recent accessibility of powerful computer vision libraries has broad applications to quantitative biology. In this thesis, I also applied computer vision techniques toward the automatic identification of DNA origami nano-barcodes. DNA nano-barcodes are a promising platform for developing next generation DNA microarray and on-chip tools. The software I developed to automatically identify and decode barcodes supplies a critical ingredient

for future developments in this platform.

Both computer vision and optogenetics are technologies undergoing rapid development. For computer vision, Moore's Law has provided ever increasing computational power at decreasing cost. Computational power alone will immediately improve both the spatial and temporal resolution of the two *C. elegans* instruments presented here. Similarly, recent trends like augmented reality for consumers and faster data connectivity standards like Intel's Thunderbolt standard, will make the low-latency high frame-rate video acquisition required for the work presented here more affordable and accessible.

In the realm of of optogenetics, the underlying technology is still nascent. The coming years will likely see the creation of a more diverse and robust optical toolkit that includes optogenetic proteins operating at new wavelengths, with greater sensitivity and narrower spectra.

Genetically encoded reporters for neural activity are also undergoing rapid development. GCaMP5, a successor to the GCaMP3 calcium indicator used here, is rumored to have better contrast, greater sensitivity and better expression levels [73]. Moreover, other labs are already racing to develop true voltage sensing proteins that can be deployed in eukaryotes [88]. This would allow researchers to visualize neural activity directly instead of inferring activity through calcium transients.

Optical neurophysiology is on its way to becoming a mainstay of neuroscience and will likely remain so for the foreseeable future. As a result, instrumentation of the sort developed here will be increasingly vital.

Where to go from here?

With the tools presented here, it is now possible to create quantitative datasets linking neural activity to behavior in *C. elegans*.

The next step will be to use these datasets to build data-driven mathematical models of neural circuits. The approach taken to study the vertebrate retina offers inspiration. The retina neuroscience community utilized decades of experimental data to build testable quantitative models of microcircuits within the retina that account for individual computations like motion detection and background motion suppression [7, 56]. In those cases, it was possible to treat collections of

neurons as linear filters with quantifiable impulse response functions. The models of the individual components could, when combined, accurately predict the larger behavior of the neural circuit.

The tools, techniques and data presented here are a first step toward toward bringing this data-driven modeling approach to *C. elegans*.

Bibliography

[1] Genome sequence of the nematode c. elegans: a platform for investigating biology. *Science (New York, N.Y.)*, 282(5396):2012–2018, December 1998. ISSN 0036-8075. URL http://www.ncbi.nlm.nih.gov/pubmed/9851916. PMID: 9851916.

[2] Faisal A Aldaye, Alison L Palmer, and Hanadi F Sleiman. Assembling materials with DNA as the guide. *Science (New York, N.Y.)*, 321(5897):1795–1799, September 2008. ISSN 1095-9203. doi: 10.1126/science.1154533. URL http://www.ncbi.nlm.nih.gov/pubmed/18818351. PMID: 18818351.

[3] Mark J Alkema, Melissa Hunter-Ensor, Niels Ringstad, and H Robert Horvitz. Tyramine functions independently of octopamine in the caenorhabditis elegans nervous system. *Neuron*, 46(2):247–260, April 2005. ISSN 0896-6273. doi: 10.1016/j.neuron.2005.02.024. URL http://www.ncbi.nlm.nih.gov/pubmed/15848803. PMID: 15848803.

[4] Z. F. Altun, L. A. Herndon, C. Crocker, R. Lints, and D. H. Hall. WormAtlas, 2002. URL http://www.wormatlas.org.

[5] Ebbe S Andersen, Mingdong Dong, Morten M Nielsen, Kasper Jahn, Ramesh Subramani, Wael Mamdouh, Monika M Golas, Bjoern Sander, Holger Stark, Cristiano L P Oliveira, Jan Skov Pedersen, Victoria Birkedal, Flemming Besenbacher, Kurt V Gothelf, and Jørgen Kjems. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature*, 459(7243):73–76, May 2009. ISSN 1476-4687. doi: 10.1038/nature07971. URL http://www.ncbi.nlm.nih.gov/pubmed/19424153. PMID: 19424153.

[6] Ryoko Ando, Hiroshi Hama, Miki Yamamoto-Hino, Hideaki Mizuno, and

Atsushi Miyawaki. An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12651–12656, October 2002. ISSN 0027-8424. doi: 10.1073/pnas.202320599. URL http://www.ncbi.nlm.nih.gov/pubmed/12271129. PMID: 12271129.

[7] Stephen A Baccus, Bence P Olveczky, Mihai Manu, and Markus Meister. A retinal circuit that computes object motion. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 28(27):6807–6817, July 2008. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.4206-07.2008. URL http://www.ncbi.nlm.nih.gov/pubmed/18596156. PMID: 18596156.

[8] Gaëtan Bellot, Mark A McClintock, Chenxiang Lin, and William M Shih. Recovery of intact DNA nanostructures after agarose gel-based separation. *Nature Methods*, 8(3):192–194, March 2011. ISSN 1548-7105. doi: 10.1038/nmeth0311-192. URL http://www.ncbi.nlm.nih.gov/pubmed/21358621. PMID: 21358621.

[9] Juliette Ben Arous, Yoshinori Tanizawa, Ithai Rabinowitch, Didier Chatenay, and William R. Schafer. Automated imaging of neuronal activity in freely behaving caenorhabditis elegans. *Journal of Neuroscience Methods*, 187(2):229–234, March 2010. ISSN 0165-0270. doi: 10.1016/j.jneumeth.2010.01.011. URL http://www.sciencedirect.com/science/article/B6T04-4Y6S7GN-3/2/f80f32c2892d34b07f1db80b870e4b53.

[10] Stefano Berri, Jordan H Boyle, Manlio Tassieri, Ian A Hope, and Netta Cohen. Forward locomotion of the nematode c. elegans is achieved through modulation of a single gait. *HFSP Journal*, 3(3):186–193, June 2009. ISSN 1955-2068. doi: 10.2976/1.3082260. URL http://www.ncbi.nlm.nih.gov/pubmed/19639043. PMID: 19639043.

[11] Eric Betzig, George H Patterson, Rachid Sougrat, O Wolf Lindwasser, Scott Olenych, Juan S Bonifacino, Michael W Davidson, Jennifer Lippincott-Schwartz, and Harald F Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science (New York, N.Y.)*, 313(5793):1642–1645,

September 2006. ISSN 1095-9203. doi: 10.1126/science.1127344. URL http://www.ncbi.nlm.nih.gov/pubmed/16902090. PMID: 16902090.

[12] Nikhil Bhatla. C. elegans - interactive neural network. http://wormweb.org/neuralnet, June 2009. URL http://wormweb.org/neuralnet.

[13] Edward Boyden. A history of optogenetics: the development of tools for controlling brain circuits with light. *F1000 Biology Reports*, 3, May 2011. ISSN 1757594X. doi: 10.3410/B3-11. URL http://f1000.com/reports/b/3/11/.

[14] Edward S Boyden, Feng Zhang, Ernst Bamberg, Georg Nagel, and Karl Deisseroth. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, 8(9):1263–1268, September 2005. ISSN 1097-6256. doi: 10.1038/nn1525. URL http://www.ncbi.nlm.nih.gov/pubmed/16116447. PMID: 16116447.

[15] Gary Bradski. The OpenCV library. *Dr. Dobb's Journal of Software Tools*, 25 (11):120, 122–125, November 2000. ISSN 1044-789X.

[16] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, Inc., 1st edition, October 2008. ISBN 0596516134.

[17] Kevin Braeckmans, Stefaan C De Smedt, Chris Roelant, Marc Leblans, Rudi Pauwels, and Joseph Demeester. Encoding microcarriers by spatial selective photobleaching. *Nature Materials*, 2(3):169–173, March 2003. ISSN 1476-1122. doi: 10.1038/nmat828. URL http://www.ncbi.nlm.nih.gov/pubmed/12612674. PMID: 12612674.

[18] John Bryden and Netta Cohen. Neural control of caenorhabditis elegans forward locomotion: the role of sensory feedback. *Biological Cybernetics*, 98(4):339–351, April 2008. ISSN 0340-1200. doi: 10.1007/s00422-008-0212-6. URL http://www.ncbi.nlm.nih.gov/pubmed/18350313. PMID: 18350313.

[19] Hieu Bui, Craig Onodera, Carson Kidwell, YerPeng Tan, Elton Graugnard, Wan Kuang, Jeunghoon Lee, William B Knowlton, Bernard Yurke, and William L Hughes. Programmable periodicity of quantum dot arrays with DNA origami nanotubes. *Nano Letters*, 10(9):3367–3372, September 2010. ISSN 1530-6992. doi: 10.1021/nl101079u. URL http://www.ncbi.nlm.nih.gov/pubmed/20681601. PMID: 20681601.

[20] Santiago Ramón y Cajal. *Advice for a Young Investigator*. The MIT Press, 1 edition, February 1999. ISBN 0262181916.

[21] M Chalfie and J Sulston. Developmental genetics of the mechanosensory neurons of caenorhabditis elegans. *Developmental Biology*, 82(2):358–370, March 1981. ISSN 0012-1606. URL http://www.ncbi.nlm.nih.gov/pubmed/7227647. PMID: 7227647.

[22] M Chalfie, J E Sulston, J G White, E Southgate, J N Thomson, and S Brenner. The neural circuit for touch sensitivity in caenorhabditis elegans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 5(4):956–64, April 1985. ISSN 0270-6474. doi: 3981252. URL http://www.ncbi.nlm.nih.gov/pubmed/3981252. PMID: 3981252.

[23] M Chalfie, Y Tu, G Euskirchen, W W Ward, and D C Prasher. Green fluorescent protein as a marker for gene expression. *Science (New York, N.Y.)*, 263(5148):802–805, February 1994. ISSN 0036-8075. URL http://www.ncbi.nlm.nih.gov/pubmed/8303295. PMID: 8303295.

[24] Beth L Chen. *Neuronal Network of C. elegans: from Anatomy to Behavior*. Thesis, Watson School of Biological Sciences, Cold Sporing Harbor Laboratory,: Cold Spring Harbor, 2007. p. 96.

[25] Beth L Chen, David H Hall, and Dmitri B Chklovskii. Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4723–4728, March 2006. ISSN 0027-8424. doi: 10.1073/pnas.0506806103. URL http://www.ncbi.nlm.nih.gov/pubmed/16537428. PMID: 16537428.

[26] Li E Cheng, Wei Song, Loren L Looger, Lily Yeh Jan, and Yuh Nung Jan. The role of the TRP channel NompC in drosophila larval and adult locomotion. *Neuron*, 67(3):373–380, August 2010. ISSN 1097-4199. doi: 10.1016/j.neuron.2010.07.004. URL http://www.ncbi.nlm.nih.gov/pubmed/20696376. PMID: 20696376.

[27] Brian Y Chow, Xue Han, Allison S Dobry, Xiaofeng Qian, Amy S Chuong, Mingjie Li, Michael A Henninger, Gabriel M Belfort, Yingxi Lin, Patrick E Monahan, and Edward S Boyden. High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature*, 463(7277): 98–102, January 2010. ISSN 1476-4687. doi: 10.1038/nature08652. URL http://www.ncbi.nlm.nih.gov/pubmed/20054397. PMID: 20054397.

[28] Nikos Chronis, Manuel Zimmer, and Cornelia I Bargmann. Microfluidics for in vivo imaging of neuronal and behavioral activity in caenorhabditis elegans. *Nature Methods*, 4(9):727–731, September 2007. ISSN 1548-7091. doi: 10.1038/nmeth1075. URL http://www.ncbi.nlm.nih.gov/pubmed/17704783. PMID: 17704783.

[29] Damon A Clark, David Biron, Piali Sengupta, and Aravinthan D T Samuel. The AFD sensory neurons encode multiple functions underlying thermotactic behavior in caenorhabditis elegans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 26(28):7444–7451, July 2006. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.1137-06.2006. URL http://www.ncbi.nlm.nih.gov/pubmed/16837592. PMID: 16837592.

[30] Damon A. Clark, Christopher V. Gabel, Harrison Gabel, and Aravinthan D. T. Samuel. Temporal activity patterns in thermosensory neurons of freely moving caenorhabditis elegans encode spatial thermal gradients. *J. Neurosci.*, 27(23):6083–6090, June 2007. doi: 10.1523/JNEUROSCI.1032-07.2007. URL http://www.jneurosci.org/cgi/content/abstract/27/23/6083.

[31] A H Cohen and P Wallén. The neuronal correlate of locomotion in fish. "Fictive swimming" induced in an in vitro preparation of the lamprey spinal cord. *Experimental Brain Research. Experimentelle Hirnforschung. Expéri-*

mentation Cérébrale, 41(1):11–18, 1980. ISSN 0014-4819. URL http://www.ncbi.nlm.nih.gov/pubmed/7461065. PMID: 7461065.

[32] NA Croll. Behavoural analysis of nematode movement. *Advances in Parasitology*, 13:71–122, 1975.

[33] NA Croll. Components and patterns in the behavior of the nematode caenorhabditis elegans. *Journal of Zoology*, 176:159–176, 1975. URL http://www.wormbase.org/db/misc/biblio?name=BEHAV%2FMOVEMENT&category=&class=Keyword&abstract=WBPaper00000075#WBPaper00000075.

[34] D F Cully, D K Vassilatis, K K Liu, P S Paress, L H Van der Ploeg, J M Schaeffer, and J P Arena. Cloning of an avermectin-sensitive glutamate-gated chloride channel from caenorhabditis elegans. *Nature*, 371(6499): 707–711, October 1994. ISSN 0028-0836. doi: 10.1038/371707a0. URL http://www.ncbi.nlm.nih.gov/pubmed/7935817. PMID: 7935817.

[35] Frédérique Cunin, Thomas A Schmedake, Jamie R Link, Yang Yang Li, Jennifer Koh, Sangeeta N Bhatia, and Michael J Sailor. Biomolecular screening with encoded porous-silicon photonic crystals. *Nature Materials*, 1(1): 39–41, September 2002. ISSN 1476-1122. doi: 10.1038/nmat702. URL http://www.ncbi.nlm.nih.gov/pubmed/12618846. PMID: 12618846.

[36] Matthew J Dejneka, Alexander Streltsov, Santona Pal, Anthony G Frutos, Christy L Powell, Kevin Yost, Po Ki Yuen, Uwe Müller, and Joydeep Lahiri. Rare earth-doped glass microbarcodes. *Proceedings of the National Academy of Sciences of the United States of America*, 100(2):389–393, January 2003. ISSN 0027-8424. doi: 10.1073/pnas.0236044100. URL http://www.ncbi.nlm.nih.gov/pubmed/12515864. PMID: 12515864.

[37] F Delcomyn. Neural basis of rhythmic behavior in animals. *Science (New York, N.Y.)*, 210(4469):492–498, October 1980. ISSN 0036-8075. URL http://www.ncbi.nlm.nih.gov/pubmed/7423199. PMID: 7423199.

[38] J A Dent, M M Smith, D K Vassilatis, and L Avery. The genetics of ivermectin resistance in caenorhabditis elegans. *Proceedings of the National*

Academy of Sciences of the United States of America, 97(6):2674–2679, March 2000. ISSN 0027-8424. URL http://www.ncbi.nlm.nih.gov/pubmed/10716995. PMID: 10716995.

[39] Hendrik Dietz, Shawn M Douglas, and William M Shih. Folding DNA into twisted and curved nanoscale shapes. *Science (New York, N.Y.)*, 325(5941):725–730, August 2009. ISSN 1095-9203. doi: 10.1126/science.1174251. URL http://www.ncbi.nlm.nih.gov/pubmed/19661424. PMID: 19661424.

[40] Shawn M Douglas, James J Chou, and William M Shih. DNA-nanotube-induced alignment of membrane proteins for NMR structure determination. *Proceedings of the National Academy of Sciences of the United States of America*, 104(16):6644–6648, April 2007. ISSN 0027-8424. doi: 10.1073/pnas.0700930104. URL http://www.ncbi.nlm.nih.gov/pubmed/17404217. PMID: 17404217.

[41] Shawn M Douglas, Hendrik Dietz, Tim Liedl, Björn Högberg, Franziska Graf, and William M Shih. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature*, 459(7245):414–418, May 2009. ISSN 1476-4687. doi: 10.1038/nature08016. URL http://www.ncbi.nlm.nih.gov/pubmed/19458720. PMID: 19458720.

[42] George Bard Ermentrout and Nancy Kopell. Frequency plateaus in a chain of weakly coupled oscillators, i. *SIAM Journal on Mathematical Analysis*, 15:215, 1984. ISSN 00361410. doi: 10.1137/0515019. URL http://link.aip.org/link/SJMAAH/v15/i2/p215/s1&Agg=doi.

[43] Behrooz Esmaeili, Jennifer M Ross, Cara Neades, 3rd Miller, David M, and Julie Ahringer. The c. elegans even-skipped homologue, vab-7, specifies DB motoneurone identity and axon trajectory. *Development (Cambridge, England)*, 129(4):853–862, February 2002. ISSN 0950-1991. URL http://www.ncbi.nlm.nih.gov/pubmed/11861469. PMID: 11861469.

[44] Clarkj Evans. The official YAML web site. http://yaml.org/, 2011. URL http://yaml.org/.

[45] Christopher Fang-Yen, Matthieu Wyart, Julie Xie, Risa Kawai, Tom Kodger, Sway Chen, Quan Wen, and Aravinthan D T Samuel. Biomechanical analysis of gait adaptation in the nematode caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (47):20323–20328, November 2010. ISSN 1091-6490. doi: 10.1073/pnas.1003016107. URL http://www.ncbi.nlm.nih.gov/pubmed/21048086. PMID: 21048086.

[46] Serge Faumont and Shawn R. Lockery. The awake behaving worm: Simultaneous imaging of neuronal activity and behavior in intact animals at millimeter scale. *J Neurophysiol*, 95(3):1976–1981, March 2006. doi: 10.1152/jn.01050.2005. URL http://jn.physiology.org/cgi/content/abstract/95/3/1976.

[47] Serge Faumont, Gary Rondeau, Tod R. Thiele, Kristy J. Lawton, Kathryn E. McCormick, Matthew Sottile, Oliver Griesbeck, Ellie S. Heckscher, William M. Roberts, Chris Q. Doe, and Shawn R. Lockery. An Image-Free Opto-Mechanical system for creating virtual environments and imaging neuronal activity in freely moving caenorhabditis elegans. *PLoS ONE*, 6(9): e24666, 2011. doi: 10.1371/journal.pone.0024666. URL http://dx.doi.org/10.1371/journal.pone.0024666.

[48] Lief Fenno, Ofer Yizhar, and Karl Deisseroth. The development and application of optogenetics. *Annual Review of Neuroscience*, 34:389–412, 2011. ISSN 1545-4126. doi: 10.1146/annurev-neuro-061010-113817. URL http://www.ncbi.nlm.nih.gov/pubmed/21692661. PMID: 21692661.

[49] Alexander P Fields and Adam E Cohen. Electrokinetic trapping at the one nanometer limit. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22):8937–8942, May 2011. ISSN 1091-6490. doi: 10.1073/pnas.1103554108. URL http://www.ncbi.nlm.nih.gov/pubmed/21562206. PMID: 21562206.

[50] Stanley Finger. *Origins of neuroscience: a history of explorations into brain function*. Oxford University Press, September 2001. ISBN 9780195146943.

[51] Sébastien Fournier-Bidoz, Travis L Jennings, Jesse M Klostranec, Winnie Fung, Alex Rhee, David Li, and Warren C W Chan. Facile and rapid one-step mass preparation of quantum-dot barcodes. *Angewandte Chemie (International Ed. in English)*, 47(30):5577–5581, 2008. ISSN 1521-3773. doi: 10.1002/anie.200800409. URL http://www.ncbi.nlm.nih.gov/pubmed/18613155. PMID: 18613155.

[52] W O Friesen and J Cang. Sensory and central mechanisms control intersegmental coordination. *Current Opinion in Neurobiology*, 11(6):678–683, December 2001. ISSN 0959-4388. URL http://www.ncbi.nlm.nih.gov/pubmed/11741017. PMID: 11741017.

[53] W O Friesen, M Poon, and G S Stent. Neuronal control of swimming in the medicinal leech. IV. identification of a network of oscillatory interneurones. *The Journal of Experimental Biology*, 75:25–43, August 1978. ISSN 0022-0949. URL http://www.ncbi.nlm.nih.gov/pubmed/702043. PMID: 702043.

[54] Arnaud Gautier, Alexandre Juillerat, Christian Heinis, Jr Corrêa, Ivan Reis, Maik Kindermann, Florent Beaufils, and Kai Johnsson. An engineered protein tag for multiprotein labeling in living cells. *Chemistry & Biology*, 15(2): 128–136, February 2008. ISSN 1074-5521. doi: 10.1016/j.chembiol.2008.01.007. URL http://www.ncbi.nlm.nih.gov/pubmed/18291317. PMID: 18291317.

[55] Gary K Geiss, Roger E Bumgarner, Brian Birditt, Timothy Dahl, Naeem Dowidar, Dwayne L Dunaway, H Perry Fell, Sean Ferree, Renee D George, Tammy Grogan, Jeffrey J James, Malini Maysuria, Jeffrey D Mitton, Paola Oliveri, Jennifer L Osborn, Tao Peng, Amber L Ratcliffe, Philippa J Webster, Eric H Davidson, Leroy Hood, and Krassen Dimitrov. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nature Biotechnology*, 26(3):317–325, March 2008. ISSN 1546-1696. doi: 10.1038/nbt1385. URL http://www.ncbi.nlm.nih.gov/pubmed/18278033. PMID: 18278033.

[56] Tim Gollisch and Markus Meister. Eye smarter than scientists believed:

neural computations in circuits of the retina. *Neuron*, 65(2):150–164, January 2010. ISSN 1097-4199. doi: 10.1016/j.neuron.2009.12.009. URL http://www.ncbi.nlm.nih.gov/pubmed/20152123. PMID: 20152123.

[57] M B Goodman, D H Hall, L Avery, and S R Lockery. Active currents regulate sensitivity and dynamic range in c. elegans neurons. *Neuron*, 20(4): 763–772, April 1998. ISSN 0896-6273. URL http://www.ncbi.nlm.nih.gov/pubmed/9581767. PMID: 9581767.

[58] P. S. Grewal and D. J. Wright. Migration of caenorhabditis elegans(Nematoda: rhabditidae) larvae towards bacteria and the nature of the bacterial stimulus. *FUND. APPL. NEMATOL.*, 15(2):159–166, 1992.

[59] Mark S Gudiksen, Lincoln J Lauhon, Jianfang Wang, David C Smith, and Charles M Lieber. Growth of nanowire superlattice structures for nanoscale photonics and electronics. *Nature*, 415(6872):617–620, February 2002. ISSN 0028-0836. doi: 10.1038/415617a. URL http://www.ncbi.nlm.nih.gov/pubmed/11832939. PMID: 11832939.

[60] Zengcai V Guo, Anne C Hart, and Sharad Ramanathan. Optical interrogation of neural circuits in caenorhabditis elegans. *Nat Meth*, advance online publication, November 2009. ISSN 1548-7105. doi: 10.1038/nmeth.1397. URL http://dx.doi.org/10.1038/nmeth.1397.

[61] Dongran Han, Suchetan Pal, Yan Liu, and Hao Yan. Folding and cutting DNA into reconfigurable topological nanostructures. *Nature Nanotechnology*, 5(10):712–717, October 2010. ISSN 1748-3395. doi: 10.1038/nnano.2010.193. URL http://www.ncbi.nlm.nih.gov/pubmed/20890274. PMID: 20890274.

[62] Dongran Han, Suchetan Pal, Jeanette Nangreave, Zhengtao Deng, Yan Liu, and Hao Yan. DNA origami with complex curvatures in three-dimensional space. *Science (New York, N.Y.)*, 332(6027):342–346, April 2011. ISSN 1095-9203. doi: 10.1126/science.1202998. URL http://www.ncbi.nlm.nih.gov/pubmed/21493857. PMID: 21493857.

[63] M Han, X Gao, J Z Su, and S Nie. Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nature Biotechnology*, 19(7):

631–635, July 2001. ISSN 1087-0156. doi: 10.1038/90228. URL http://www.ncbi.nlm.nih.gov/pubmed/11433273. PMID: 11433273.

[64] Xue Han and Edward S. Boyden. Multiple-Color optical activation, silencing, and desynchronization of neural activity, with Single-Spike temporal resolution. *PLoS ONE*, 2(3):e299, 2007. doi: 10.1371/journal.pone.0000299. URL http://dx.plos.org/10.1371/journal.pone.0000299.

[65] Todd W Harris, Igor Antoshechkin, Tamberlyn Bieri, Darin Blasiar, Juancarlos Chan, Wen J Chen, Norie De La Cruz, Paul Davis, Margaret Duesbury, Ruihua Fang, Jolene Fernandes, Michael Han, Ranjana Kishore, Raymond Lee, Hans-Michael Müller, Cecilia Nakamura, Philip Ozersky, Andrei Petcherski, Arun Rangarajan, Anthony Rogers, Gary Schindelman, Erich M Schwarz, Mary Ann Tuli, Kimberly Van Auken, Daniel Wang, Xiaodong Wang, Gary Williams, Karen Yook, Richard Durbin, Lincoln D Stein, John Spieth, and Paul W Sternberg. WormBase: a comprehensive resource for nematode research. *Nucleic Acids Research*, 38(Database issue):D463–467, January 2010. ISSN 1362-4962. doi: 10.1093/nar/gkp952. URL http://www.ncbi.nlm.nih.gov/pubmed/19910365. PMID: 19910365.

[66] A. C. Hart. Behavior. In The C. elegans Research Community, editor, *WormBook*, number 82. WormBook, the c. elegans research community edition, July 2006. ISBN 1551-507. URL http://www.wormbook.org/chapters/www_behavior/behavior.html.

[67] Gal Haspel, Michael J O'Donovan, and Anne C Hart. Motoneurons dedicated to either forward or backward locomotion in the nematode caenorhabditis elegans. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(33):11151–11156, August 2010. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.2244-10.2010. URL http://www.ncbi.nlm.nih.gov/pubmed/20720122. PMID: 20720122.

[68] Mike Heilemann, Sebastian van de Linde, Mark Schüttpelz, Robert Kasper, Britta Seefeldt, Anindita Mukherjee, Philip Tinnefeld, and Markus Sauer. Subdiffraction-resolution fluorescence imaging with conventional fluores-

cent probes. *Angewandte Chemie (International Ed. in English)*, 47(33): 6172–6176, 2008. ISSN 1521-3773. doi: 10.1002/anie.200802376. URL http://www.ncbi.nlm.nih.gov/pubmed/18646237. PMID: 18646237.

[69] Stefan W Hell. Far-field optical nanoscopy. *Science (New York, N.Y.)*, 316(5828):1153–1158, May 2007. ISSN 1095-9203. doi: 10.1126/science.1137395. URL http://www.ncbi.nlm.nih.gov/pubmed/17525330. PMID: 17525330.

[70] Stefan W Hell. Microscopy and its focal switch. *Nature Methods*, 6(1):24–32, January 2009. ISSN 1548-7105. doi: 10.1038/nmeth.1291. URL http://www.ncbi.nlm.nih.gov/pubmed/19116611. PMID: 19116611.

[71] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19(11):780–782, June 1994. doi: 10.1364/OL.19.000780. URL http://ol.osa.org/abstract.cfm?URI=ol-19-11-780.

[72] Carl W Helstrom. *Statistical Theory of Signal Detection*. International series of monographs in electronics and instrumentation, v. 9. Pergamon Press, Oxford, New York, 2d ed., rev. and enl edition, 1968. ISBN 0080132650.

[73] Andrew Hires. GCaMP5 is out, November 2011. URL http://brainwindows.wordpress.com/2011/11/15/gcamp5-is-out/.

[74] Bo Huang, Hazen Babcock, and Xiaowei Zhuang. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell*, 143(7):1047–1058, December 2010. ISSN 1097-4172. doi: 10.1016/j.cell.2010.12.002. URL http://www.ncbi.nlm.nih.gov/pubmed/21168201. PMID: 21168201.

[75] Cynthia L Hughes and John B Thomas. A sensory feedback circuit coordinates muscle activity in drosophila. *Molecular and Cellular Neurosciences*, 35(2):383–396, June 2007. ISSN 1044-7431. doi: 10.1016/j.mcn.2007.04.001. URL http://www.ncbi.nlm.nih.gov/pubmed/17498969. PMID: 17498969.

[76] Yuichi Iino and Kazushi Yoshida. Parallel use of two behavioral mechanisms for chemotaxis in caenorhabditis elegans. *The Journal of Neuroscience:*

The Official Journal of the Society for Neuroscience, 29(17):5370–5380, April 2009. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.3633-08.2009. URL http://www.ncbi.nlm.nih.gov/pubmed/19403805. PMID: 19403805.

[77] E. T Jaynes. *Probability theory : the logic of science*. Cambridge University Press, Cambridge, UK, 2003. ISBN 0521592712 (hardback).

[78] Sara A Jones, Sang-Hee Shim, Jiang He, and Xiaowei Zhuang. Fast, three-dimensional super-resolution imaging of live cells. *Nature Methods*, 8(6): 499–508, June 2011. ISSN 1548-7105. doi: 10.1038/nmeth.1605. URL http://www.ncbi.nlm.nih.gov/pubmed/21552254. PMID: 21552254.

[79] Ralf Jungmann, Christian Steinhauer, Max Scheible, Anton Kuzyk, Philip Tinnefeld, and Friedrich C Simmel. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Letters*, 10(11):4756–4761, November 2010. ISSN 1530-6992. doi: 10.1021/nl103427w. URL http://www.ncbi.nlm.nih.gov/pubmed/20957983. PMID: 20957983.

[80] Eric R. Kandel. *In Search of Memory: The Emergence of a New Science of Mind*. W. W. Norton & Company, 1 edition, March 2007. ISBN 0393329372.

[81] Eric R. Kandel, James Harris Schwartz, and Thomas M. Jessell. *Essentials of neural science and behavior*. McGraw-Hill Professional, 1995. ISBN 9780838522455.

[82] Jan Karbowski, Gary Schindelman, Christopher J Cronin, Adeline Seah, and Paul W Sternberg. Systems level circuit model of c. elegans undulatory locomotion: mathematical modeling and molecular genetics. *Journal of Computational Neuroscience*, 24(3):253–276, June 2008. ISSN 1573-6873. doi: 10.1007/s10827-007-0054-6. URL http://www.ncbi.nlm.nih.gov/pubmed/17768672. PMID: 17768672.

[83] Taizo Kawano, Michelle D Po, Shangbang Gao, George Leung, William S Ryu, and Mei Zhen. An imbalancing act: Gap junctions reduce the backward motor circuit activity to bias c. elegans for forward locomotion. *Neuron*, 72 (4):572–586, November 2011. ISSN 1097-4199. doi: 10.1016/j.neuron.

2011.09.005. URL http://www.ncbi.nlm.nih.gov/pubmed/22099460. PMID: 22099460.

[84] Yonggang Ke, Shawn M Douglas, Minghui Liu, Jaswinder Sharma, Anchi Cheng, Albert Leung, Yan Liu, William M Shih, and Hao Yan. Multilayer DNA origami packed on a square lattice. *Journal of the American Chemical Society*, 131(43):15903–15908, November 2009. ISSN 1520-5126. doi: 10.1021/ja906381y. URL http://www.ncbi.nlm.nih.gov/pubmed/19807088. PMID: 19807088.

[85] Antje Keppler, Susanne Gendreizig, Thomas Gronemeyer, Horst Pick, Horst Vogel, and Kai Johnsson. A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nature Biotechnology*, 21(1): 86–89, January 2003. ISSN 1087-0156. doi: 10.1038/nbt765. URL http://www.ncbi.nlm.nih.gov/pubmed/12469133. PMID: 12469133.

[86] Kei-ichiro Kitamura, Shigetoyo Amano, and Ryuji Hosono. Contribution of neurons to habituation to mechanical stimulation in caenorhabditis elegans. *Journal of Neurobiology*, 46(1):29–40, 2001. doi: 10.1002/1097-4695(200101)46:1<29::AID-NEU3>3.0.CO;2-8. URL http://dx.doi.org/10.1002/1097-4695(200101)46:1%3C29::AID-NEU3%3E3.0.CO;2-8.

[87] Teresa Klein, Anna Löschberger, Sven Proppert, Steve Wolter, Sebastian van de Linde, and Markus Sauer. Live-cell dSTORM with SNAP-tag fusion proteins. *Nature Methods*, 8(1):7–9, January 2011. ISSN 1548-7105. doi: 10.1038/nmeth0111-7b. URL http://www.ncbi.nlm.nih.gov/pubmed/21191367. PMID: 21191367.

[88] Joel M Kralj, Daniel R Hochbaum, Adam D Douglass, and Adam E Cohen. Electrical spiking in escherichia coli probed with a fluorescent voltage-indicating protein. *Science (New York, N.Y.)*, 333(6040):345–348, July 2011. ISSN 1095-9203. doi: 10.1126/science.1204763. URL http://www.ncbi.nlm.nih.gov/pubmed/21764748. PMID: 21764748.

[89] Jr Kristan, W B and R L Calabrese. Rhythmic swimming activity in neurones of the isolated nerve cord of the leech. *The Journal of Experimental*

Biology, 65(3):643–668, December 1976. ISSN 0022-0949. URL http://www.ncbi.nlm.nih.gov/pubmed/1018167. PMID: 1018167.

[90] Biji T Kurien and R Hal Scofield. Extraction of nucleic acid fragments from gels. *Analytical Biochemistry*, 302(1):1–9, March 2002. ISSN 0003-2697. doi: 10.1006/abio.2001.5526. URL http://www.ncbi.nlm.nih.gov/pubmed/11846370. PMID: 11846370.

[91] Hidehito Kuroyanagi, Tetsuo Kobayashi, Shohei Mitani, and Masatoshi Hagiwara. Transgenic alternative-splicing reporters reveal tissue-specific expression profiles and regulation mechanisms in vivo. *Nature Methods*, 3(11):909–915, November 2006. ISSN 1548-7091. doi: 10.1038/nmeth944. URL http://www.ncbi.nlm.nih.gov/pubmed/17060915. PMID: 17060915.

[92] Andrew M Leifer, Christopher Fang-Yen, Marc Gershow, Mark J Alkema, and Aravinthan D T Samuel. Optogenetic manipulation of neural activity in freely moving caenorhabditis elegans. *Nature Methods*, 8(2):147–152, February 2011. ISSN 1548-7105. doi: 10.1038/nmeth.1554. URL http://www.ncbi.nlm.nih.gov/pubmed/21240279. PMID: 21240279.

[93] Jeffrey M Levsky, Shailesh M Shenoy, Rossanna C Pezo, and Robert H Singer. Single-cell gene expression profiling. *Science (New York, N.Y.)*, 297(5582):836–840, August 2002. ISSN 1095-9203. doi: 10.1126/science.1072241. URL http://www.ncbi.nlm.nih.gov/pubmed/12161654. PMID: 12161654.

[94] Wei Li, Zhaoyang Feng, Paul W Sternberg, and X Z Shawn Xu. A c. elegans stretch receptor neuron revealed by a mechanosensitive TRP channel homologue. *Nature*, 440(7084):684–687, March 2006. ISSN 1476-4687. doi: 10.1038/nature04538. URL http://www.ncbi.nlm.nih.gov/pubmed/16572173. PMID: 16572173.

[95] Xiao Li, Tieqiang Wang, Junhu Zhang, Difu Zhu, Xun Zhang, Yang Ning, Hao Zhang, and Bai Yang. Controlled fabrication of fluorescent barcode nanorods. *ACS Nano*, 4(8):4350–4360, August 2010. ISSN 1936-086X.

doi: 10.1021/nn9017137. URL http://www.ncbi.nlm.nih.gov/pubmed/20731421. PMID: 20731421.

[96] Yougen Li, Yen Thi Hong Cu, and Dan Luo. Multiplexed detection of pathogen DNA with DNA-based fluorescence nanobarcodes. *Nature Biotechnology*, 23(7):885–889, July 2005. ISSN 1087-0156. doi: 10.1038/nbt1106. URL http://www.ncbi.nlm.nih.gov/pubmed/15951805. PMID: 15951805.

[97] Tim Liedl, Björn Högberg, Jessica Tytell, Donald E Ingber, and William M Shih. Self-assembly of three-dimensional prestressed tensegrity structures from DNA. *Nature Nanotechnology*, 5(7):520–524, July 2010. ISSN 1748-3395. doi: 10.1038/nnano.2010.107. URL http://www.ncbi.nlm.nih.gov/pubmed/20562873. PMID: 20562873.

[98] Jana F Liewald, Martin Brauner, Greg J Stephens, Magali Bouhours, Christian Schultheis, Mei Zhen, and Alexander Gottschalk. Optogenetic analysis of synaptic function. *Nat Meth*, 5(10):895–902, October 2008. ISSN 1548-7091. doi: 10.1038/nmeth.1252. URL http://dx.doi.org/10.1038/nmeth.1252.

[99] Chenxiang Lin, Yan Liu, and Hao Yan. Self-assembled combinatorial encoding nanoarrays for multiplexed biosensing. *Nano Letters*, 7(2):507–512, February 2007. ISSN 1530-6984. doi: 10.1021/nl062998n. URL http://www.ncbi.nlm.nih.gov/pubmed/17298017. PMID: 17298017.

[100] Chenxiang Lin, Yan Liu, and Hao Yan. Designer DNA nanoarchitectures. *Biochemistry*, 48(8):1663–1674, March 2009. ISSN 1520-4995. doi: 10.1021/bi802324w. URL http://www.ncbi.nlm.nih.gov/pubmed/19199428. PMID: 19199428.

[101] Chenxiang Lin, Ralf Jungmann, Andrew M Leifer, Chao Li, Daniel Levner, William M Shih, and Peng Yin. Sub-micrometer geometrically encoded fluorescent barcodes Self-Assembled from DNA. Submitted to Nature Nanotechnology, October 2011.

[102] J T Lis and R Schleif. Size fractionation of double-stranded DNA by precipitation with polyethylene glycol. *Nucleic Acids Research*, 2(3):383–389,

March 1975. ISSN 0305-1048. URL http://www.ncbi.nlm.nih.gov/pubmed/236548. PMID: 236548.

[103] Qiang Liu, Bojun Chen, Eric Gaier, Jaya Joshi, and Zhao-Wen Wang. Low conductance gap junctions mediate specific electrical coupling in body-wall muscle cells of caenorhabditis elegans. *The Journal of Biological Chemistry*, 281(12):7881–7889, March 2006. ISSN 0021-9258. doi: 10.1074/jbc.M512382200. URL http://www.ncbi.nlm.nih.gov/pubmed/16434400. PMID: 16434400.

[104] Wenyan Liu, Hong Zhong, Risheng Wang, and Nadrian C Seeman. Crystalline two-dimensional DNA-origami arrays. *Angewandte Chemie (International Ed. in English)*, 50(1):264–267, January 2011. ISSN 1521-3773. doi: 10.1002/anie.201005911. URL http://www.ncbi.nlm.nih.gov/pubmed/21053236. PMID: 21053236.

[105] Jean Livet, Tamily A Weissman, Hyuno Kang, Ryan W Draft, Ju Lu, Robyn A Bennis, Joshua R Sanes, and Jeff W Lichtman. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*, 450(7166):56–62, November 2007. ISSN 1476-4687. doi: 10.1038/nature06293. URL http://www.ncbi.nlm.nih.gov/pubmed/17972876. PMID: 17972876.

[106] Bill Livolsi. Projectors and PC gaming. *ProjectorCentral.com*, August 2008. URL http://web.archive.org/web/20110104033820/http://www.projectorcentral.com/pc_gaming_projectors.htm.

[107] S. R. Lockery, K. J. Lawton, J. C. Doll, S. Faumont, S. M. Coulthard, T. R. Thiele, N. Chronis, K. E. McCormick, M. B. Goodman, and B. L. Pruitt. Artificial dirt: Microfluidic substrates for nematode neurobiology and behavior. *J Neurophysiol*, 99(6):3136–3143, June 2008. doi: 10.1152/jn.91327.2007. URL http://jn.physiology.org/cgi/content/abstract/99/6/3136.

[108] Jon Loeliger. *Version Control with Git*. O'Reilly Media, Inc., May 2009. ISBN 9780596520120.

[109] Michael A Long and Michale S Fee. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*, 456(7219):189–194, November 2008. ISSN 1476-4687. doi: 10.1038/nature07448. URL http://www.ncbi.nlm.nih.gov/pubmed/19005546. PMID: 19005546.

[110] Kyle Lund, Anthony J Manzo, Nadine Dabby, Nicole Michelotti, Alexander Johnson-Buck, Jeanette Nangreave, Steven Taylor, Renjun Pei, Milan N Stojanovic, Nils G Walter, Erik Winfree, and Hao Yan. Molecular robots guided by prescriptive landscapes. *Nature*, 465(7295):206–210, May 2010. ISSN 1476-4687. doi: 10.1038/nature09012. URL http://www.ncbi.nlm.nih.gov/pubmed/20463735. PMID: 20463735.

[111] Marco Mank and Oliver Griesbeck. Genetically encoded calcium indicators. *Chemical Reviews*, 108(5):1550–1564, May 2008. ISSN 1520-6890. doi: 10.1021/cr078213v. URL http://www.ncbi.nlm.nih.gov/pubmed/18447377. PMID: 18447377.

[112] Lionel Marcon, Bronwyn J Battersby, Andreas Rühmann, Kym Ford, Matthew Daley, Gwendolyn A Lawrie, and Matt Trau. 'On-the-fly' optical encoding of combinatorial peptide libraries for profiling of protease specificity. *Molecular bioSystems*, 6(1):225–233, January 2010. ISSN 1742-2051. doi: 10.1039/b909087h. URL http://www.ncbi.nlm.nih.gov/pubmed/20024084. PMID: 20024084.

[113] E Marder and R L Calabrese. Principles of rhythmic motor pattern generation. *Physiological Reviews*, 76(3):687–717, July 1996. ISSN 0031-9333. URL http://www.ncbi.nlm.nih.gov/pubmed/8757786. PMID: 8757786.

[114] MATLAB. *Version 7.11.0 (R2010b)*. The MathWorks Inc., 2010.

[115] A Miyawaki, J Llopis, R Heim, J M McCaffery, J A Adams, M Ikura, and R Y Tsien. Fluorescent indicators for ca2+ based on green fluorescent proteins and calmodulin. *Nature*, 388(6645):882–887, August 1997. ISSN 0028-0836. doi: 10.1038/42264. URL http://www.ncbi.nlm.nih.gov/pubmed/9278050. PMID: 9278050.

[116] Atsushi Miyawaki. Innovations in the imaging of brain functions using fluorescent proteins. *Neuron*, 48(2):189–199, October 2005. ISSN 0896-6273. doi: 10.1016/j.neuron.2005.10.003. URL http://www.ncbi.nlm.nih.gov/pubmed/16242400. PMID: 16242400.

[117] Georg Nagel, Tanjef Szellas, Wolfram Huhn, Suneel Kateriya, Nona Adeishvili, Peter Berthold, Doris Ollig, Peter Hegemann, and Ernst Bamberg. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proceedings of the National Academy of Sciences of the United States of America*, 100(24):13940 –13945, November 2003. doi: 10.1073/pnas.1936192100. URL http://www.pnas.org/content/100/24/13940.abstract.

[118] Georg Nagel, Martin Brauner, Jana F. Liewald, Nona Adeishvili, Ernst Bamberg, and Alexander Gottschalk. Light activation of channelrhodopsin-2 in excitable cells of caenorhabditis elegans triggers rapid behavioral responses. *Current Biology*, 15(24):2279–2284, December 2005. ISSN 0960-9822. doi: 10.1016/j.cub.2005.11.032. URL http://www.sciencedirect.com/science/article/B6VRT-4HV7CDJ-W/2/c2d98908e8dbee2808aa298cf3c7e0d2.

[119] Jeanette Nangreave, Dongran Han, Yan Liu, and Hao Yan. DNA origami: a history and current perspective. *Current Opinion in Chemical Biology*, 14 (5):608–615, October 2010. ISSN 1879-0402. doi: 10.1016/j.cbpa.2010.06.182. URL http://www.ncbi.nlm.nih.gov/pubmed/20643573. PMID: 20643573.

[120] S R Nicewarner-Pena, R G Freeman, B D Reiss, L He, D J Pena, I D Walton, R Cromer, C D Keating, and M J Natan. Submicrometer metallic barcodes. *Science (New York, N.Y.)*, 294(5540):137–141, October 2001. ISSN 0036-8075. doi: 10.1126/science.294.5540.137. URL http://www.ncbi.nlm.nih.gov/pubmed/11588257. PMID: 11588257.

[121] E Niebur and P Erdös. Theory of the locomotion of nematodes: Dynamics of undulatory progression on a surface. *Biophysical Journal*, 60(5):1132–1146,

November 1991. ISSN 0006-3495. URL http://www.ncbi.nlm.nih.gov/pubmed/19431807. PMID: 19431807.

[122] Alan V Oppenheim and Ronald W Schafer. *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, N.J, 1975. ISBN 0132146355.

[123] The Video Lan Organization. VideoLAN - official page for VLC media player, the open source video framework! http://www.videolan.org, 2012. URL http://www.videolan.org.

[124] Suchetan Pal, Zhengtao Deng, Baoquan Ding, Hao Yan, and Yan Liu. DNA-origami-directed self-assembly of discrete silver-nanoparticle architectures. *Angewandte Chemie (International Ed. in English)*, 49(15):2700–2704, April 2010. ISSN 1521-3773. doi: 10.1002/anie.201000330. URL http://www.ncbi.nlm.nih.gov/pubmed/20235262. PMID: 20235262.

[125] R A Pearce and W O Friesen. Intersegmental coordination of leech swimming: comparison of in situ and isolated nerve cord activity with body wall movement. *Brain Research*, 299(2):363–366, May 1984. ISSN 0006-8993. URL http://www.ncbi.nlm.nih.gov/pubmed/6733455. PMID: 6733455.

[126] Beverly J Piggott, Jie Liu, Zhaoyang Feng, Seth A Wescott, and X Z Shawn Xu. The neural circuits and synaptic mechanisms underlying motor initiation in c. elegans. *Cell*, 147(4):922–933, November 2011. ISSN 1097-4172. doi: 10.1016/j.cell.2011.08.053. URL http://www.ncbi.nlm.nih.gov/pubmed/22078887. PMID: 22078887.

[127] Jennifer K. Pirri, Adam D. McPherson, Jamie L. Donnelly, Michael M. Francis, and Mark J. Alkema. A tyramine-gated chloride channel coordinates distinct motor programs of a caenorhabditis elegans escape response. *Neuron*, 62(4):526–538, May 2009. ISSN 0896-6273. doi: 10.1016/j.neuron.2009.04.013. PMID: 19477154 PMCID: 2804440.

[128] Daniel C Pregibon, Mehmet Toner, and Patrick S Doyle. Multifunctional encoded particles for high-throughput biomolecule analysis. *Science (New York, N.Y.)*, 315(5817):1393–1396, March 2007. ISSN 1095-9203. doi: 10.

1126/science.1134929. URL http://www.ncbi.nlm.nih.gov/pubmed/17347435. PMID: 17347435.

[129] Arivazhagan Rajendran, Masayuki Endo, Yousuke Katsuda, Kumi Hidaka, and Hiroshi Sugiyama. Photo-cross-linking-assisted thermal stability of DNA origami structures and its application for higher-temperature self-assembly. *Journal of the American Chemical Society*, 133(37):14488–14491, September 2011. ISSN 1520-5126. doi: 10.1021/ja204546h. URL http://www.ncbi.nlm.nih.gov/pubmed/21859143. PMID: 21859143.

[130] W. S. Rasband. *Image J*. National Institutes of Health, Bethesda, Maryland, USA, 1997. URL http://imagej.nih.gov/ij/.

[131] F. Reif. *Fundamentals of statistical and thermal physics*. McGraw-Hill series in fundamentals of physics. McGraw-Hill, New York, 1965.

[132] J E Richmond, W S Davis, and E M Jorgensen. UNC-13 is required for synaptic vesicle fusion in c. elegans. *Nature Neuroscience*, 2(11):959–964, November 1999. ISSN 1097-6256. doi: 10.1038/14755. URL http://www.ncbi.nlm.nih.gov/pubmed/10526333. PMID: 10526333.

[133] Niels Ringstad and H Robert Horvitz. FMRFamide neuropeptides and acetylcholine synergistically inhibit egg-laying by c. elegans. *Nature Neuroscience*, 11(10):1168–1176, October 2008. ISSN 1546-1726. doi: 10.1038/nn.2186. URL http://www.ncbi.nlm.nih.gov/pubmed/18806786. PMID: 18806786.

[134] A Roghani, J Feldman, S A Kohan, A Shirzadi, C B Gundersen, N Brecha, and R H Edwards. Molecular cloning of a putative vesicular transporter for acetylcholine. *Proceedings of the National Academy of Sciences of the United States of America*, 91(22):10620–10624, October 1994. ISSN 0027-8424. URL http://www.ncbi.nlm.nih.gov/pubmed/7938002. PMID: 7938002.

[135] Paul W K Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, March 2006. ISSN 1476-4687. doi: 10.1038/nature04586. URL http://www.ncbi.nlm.nih.gov/pubmed/16541064. PMID: 16541064.

[136] Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–795, October 2006. ISSN 1548-7091. doi: 10.1038/nmeth929. URL http://www.ncbi.nlm.nih.gov/pubmed/16896339. PMID: 16896339.

[137] William S. Ryu and Aravinthan D. T. Samuel. Thermotaxis in caenorhabditis elegans analyzed by measuring responses to defined thermal stimuli. *Jounral of Neuroscience*, 22(13):5727–5733, July 2002. doi: 20026542. URL http://www.jneurosci.org/cgi/content/abstract/22/13/5727.

[138] William Schafer. Neurophysiological methods in c. elegans: an introduction. *WormBook*, 2006. ISSN 15518507. doi: 10.1895/wormbook.1.111.1. URL http://www.wormbook.org/chapters/www_intromethodsneurophys/intromethodsneurophys.html.

[139] N C Seeman. Nucleic acid junctions and lattices. *Journal of Theoretical Biology*, 99(2):237–247, November 1982. ISSN 0022-5193. URL http://www.ncbi.nlm.nih.gov/pubmed/6188926. PMID: 6188926.

[140] Alexey Sharonov and Robin M Hochstrasser. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proceedings of the National Academy of Sciences of the United States of America*, 103(50):18911–18916, December 2006. ISSN 0027-8424. doi: 10.1073/pnas.0609643104. URL http://www.ncbi.nlm.nih.gov/pubmed/17142314. PMID: 17142314.

[141] William M Shih and Chenxiang Lin. Knitting complex weaves with DNA origami. *Current Opinion in Structural Biology*, 20(3):276–282, June 2010. ISSN 1879-033X. doi: 10.1016/j.sbi.2010.03.009. URL http://www.ncbi.nlm.nih.gov/pubmed/20456942. PMID: 20456942.

[142] Wei Song, Maika Onishi, Lily Yeh Jan, and Yuh Nung Jan. Peripheral multidendritic sensory neurons are necessary for rhythmic locomotion behavior in drosophila larvae. *Proceedings of the National Academy of Sciences of the United States of America*, 104(12):5199–5204, March 2007. ISSN 0027-

8424. doi: 10.1073/pnas.0700895104. URL http://www.ncbi.nlm.nih.gov/pubmed/17360325. PMID: 17360325.

[143] David Michael Stavens. *Learning to drive perception for autonomous cars*. Ph.D. dissertation, Stanford University, 2011.

[144] Christian Steinhauer, Carsten Forthmann, Jan Vogelsang, and Philip Tinnefeld. Superresolution microscopy on the basis of engineered dark states. *Journal of the American Chemical Society*, 130(50):16840–16841, December 2008. ISSN 1520-5126. doi: 10.1021/ja806590m. URL http://www.ncbi.nlm.nih.gov/pubmed/19053449. PMID: 19053449.

[145] Christian Steinhauer, Ralf Jungmann, Thomas L Sobey, Friedrich C Simmel, and Philip Tinnefeld. DNA origami as a nanoscopic ruler for super-resolution microscopy. *Angewandte Chemie (International Ed. in English)*, 48(47):8870–8873, 2009. ISSN 1521-3773. doi: 10.1002/anie.200903308. URL http://www.ncbi.nlm.nih.gov/pubmed/19830751. PMID: 19830751.

[146] Jeffrey N Stirman, Martin Brauner, Alexander Gottschalk, and Hang Lu. High-throughput study of synaptic transmission at the neuromuscular junction enabled by optogenetics and microfluidics. *Journal of Neuroscience Methods*, 191(1):90–93, August 2010. ISSN 1872-678X. doi: 10.1016/j.jneumeth.2010.05.019. URL http://www.ncbi.nlm.nih.gov/pubmed/20538016. PMID: 20538016.

[147] Jeffrey N Stirman, Matthew M Crane, Steven J Husson, Sebastian Wabnig, Christian Schultheis, Alexander Gottschalk, and Hang Lu. Real-time multimodal optical control of neurons and muscles in freely behaving caenorhabditis elegans. *Nature methods*, 8(2):153–158, February 2011. ISSN 1548-7105. doi: 10.1038/nmeth.1555. URL http://www.ncbi.nlm.nih.gov/pubmed/21240278. PMID: 21240278.

[148] Jeffrey N Stirman, Matthew M Crane, Steven J Husson, Alexander Gottschalk, and Hang Lu. A multispectral optical illumination system with precise spatiotemporal control for the manipulation of optogenetic

reagents. *Nature protocols*, 7(2):207–220, February 2012. ISSN 1750-2799. doi: 10.1038/nprot.2011.433. URL http://www.ncbi.nlm.nih.gov/pubmed/22240583. PMID: 22240583.

[149] J Sulston, Z Du, K Thomas, R Wilson, L Hillier, R Staden, N Halloran, P Green, J Thierry-Mieg, and L Qiu. The c. elegans genome sequencing project: a beginning. *Nature*, 356(6364):37–41, March 1992. ISSN 0028-0836. doi: 10.1038/356037a0. URL http://www.ncbi.nlm.nih.gov/pubmed/1538779. PMID: 1538779.

[150] Stephanie Szobota, Pau Gorostiza, Filippo Del Bene, Claire Wyart, Doris L Fortin, Kathleen D Kolstad, Orapim Tulyathan, Matthew Volgraf, Rika Numano, Holly L Aaron, Ethan K Scott, Richard H Kramer, John Flannery, Herwig Baier, Dirk Trauner, and Ehud Y Isacoff. Remote control of neuronal activity with a light-gated glutamate receptor. *Neuron*, 54(4):535–545, May 2007. ISSN 0896-6273. doi: 10.1016/j.neuron.2007.05.010. URL http://www.ncbi.nlm.nih.gov/pubmed/17521567. PMID: 17521567.

[151] The C. Elegans Research Community, editor. *WormBook*. 2011. URL http://www.wormbook.org.

[152] Sebastian Thrun, Mike Montemerlo, Hendrik Dahlkamp, David Stavens, Andrei Aron, James Diebel, Philip Fong, John Gale, Morgan Halpenny, Gabriel Hoffmann, Kenny Lau, Celia Oakley, Mark Palatucci, Vaughan Pratt, Pascal Stang, Sven Strohband, Cedric Dupont, Lars-Erik Jendrossek, Christian Koelen, Charles Markey, Carlo Rummel, Joe Niekerk, Eric Jensen, Philippe Alessandrini, Gary Bradski, Bob Davies, Scott Ettinger, Adrian Kaehler, Ara Nefian, and Pamela Mahoney. Stanley: The robot that won the DARPA grand challenge. In Martin Buehler, Karl Iagnemma, and Sanjiv Singh, editors, *The 2005 DARPA Grand Challenge*, volume 36, pages 1–43. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-73428-4, 978-3-540-73429-1. URL http://www.springerlink.com/content/ro1240114858137n/.

[153] Lin Tian, S Andrew Hires, Tianyi Mao, Daniel Huber, M Eugenia Chiappe, Sreekanth H Chalasani, Leopoldo Petreanu, Jasper Akerboom, Sean A McK-

inney, Eric R Schreiter, Cornelia I Bargmann, Vivek Jayaraman, Karel Svoboda, and Loren L Looger. Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods*, 6(12):875–881, December 2009. ISSN 1548-7105. doi: 10.1038/nmeth.1398. URL http://www.ncbi.nlm.nih.gov/pubmed/19898485. PMID: 19898485.

[154] ViaLUX. ViALUX messtechnik + bildverarbeitung GmbH, products, DLP® technology, ALP controller. http://www.vialux.de/HTML/en_alpcontr.htm, 2009. URL http://www.vialux.de/HTML/en_alpcontr.htm.

[155] Jan Vogelsang, Christian Steinhauer, Carsten Forthmann, Ingo H Stein, Britta Person-Skegro, Thorben Cordes, and Philip Tinnefeld. Make them blink: Probes for Super-Resolution microscopy. *ChemPhysChem*, 11 (12):2475–2490, August 2010. ISSN 1439-7641. doi: 10.1002/cphc.201000189. URL http://onlinelibrary.wiley.com/doi/10.1002/cphc.201000189/abstract.

[156] Stephen E Von Stetina, Millet Treinin, and David M Miller. The motor circuit. *International Review of Neurobiology*, 69:125–167, 2006. ISSN 0074-7742. doi: 10.1016/S0074-7742(05)69005-8. URL http://www.ncbi.nlm.nih.gov/pubmed/16492464. PMID: 16492464.

[157] L. A Wainstein and V. D Zubakov. *Extraction of signals from noise*. Prentice-Hall, Englewood Cliffs, N. J., 1962.

[158] Nils G Walter, Cheng-Yen Huang, Anthony J Manzo, and Mohamed A Sobhy. Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nature Methods*, 5(6):475–489, June 2008. ISSN 1548-7105. doi: 10.1038/nmeth.1215. URL http://www.ncbi.nlm.nih.gov/pubmed/18511916. PMID: 18511916.

[159] Quan Wen, Elizabeth Hulme, Sway Chen, Xinyu Liu, Marc Gershow, Andrew M Leifer, Victoria Butler, Christopher Fang-Yen, William R Schafer, George Whitesides, Matthieu Wyart, Dmitri B Chklovskii, and Aravinthan D T Samuel. Bending waves during caenorhabditis elegans locomotion are driven and organized by proprioceptive coupling. *Nature (submitted)*, 2011.

[160] J. G. White, Eileen Southgate, J. N. Thomson, and S. Brenner. The structure of the ventral nerve cord of caenorhabditis elegans. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 275(938): 327 –348, 1976. doi: 10.1098/rstb.1976.0086. URL http://rstb.royalsocietypublishing.org/content/275/938/327.abstract.

[161] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner. The structure of the nervous system of the nematode caenorhabditis elegans. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 314(1165):1–340, November 1986. ISSN 00804622. URL http://www.jstor.org/stable/2990196.

[162] S R Wicks, C J Roehrig, and C H Rankin. A dynamic network simulation of the nematode tap withdrawal circuit: predictions concerning synaptic function using behavioral criteria. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 16(12):4017–4031, June 1996. ISSN 0270-6474. URL http://www.ncbi.nlm.nih.gov/pubmed/8656295. PMID: 8656295.

[163] Sungwook Woo and Paul W K Rothemund. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nature Chemistry*, 3(8):620–627, August 2011. ISSN 1755-4349. doi: 10.1038/nchem.1070. URL http://www.ncbi.nlm.nih.gov/pubmed/21778982. PMID: 21778982.

[164] Claire Wyart, Filippo Del Bene, Erica Warp, Ethan K Scott, Dirk Trauner, Herwig Baier, and Ehud Y Isacoff. Optogenetic dissection of a behavioural module in the vertebrate spinal cord. *Nature*, 461(7262):407–410, September 2009. ISSN 1476-4687. doi: 10.1038/nature08323. URL http://www.ncbi.nlm.nih.gov/pubmed/19759620. PMID: 19759620.

[165] Ming Xiao, Eunice Wan, Catherine Chu, Wen-Chi Hsueh, Yang Cao, and Pui-Yan Kwok. Direct determination of haplotypes from single DNA molecules. *Nature Methods*, 6(3):199–201, March 2009. ISSN 1548-7105. doi: 10.1038/nmeth.1301. URL http://www.ncbi.nlm.nih.gov/pubmed/19198595. PMID: 19198595.

[166] Hongxia Xu, Michael Y Sha, Edith Y Wong, Janet Uphoff, Yanzhang Xu, Joseph A Treadway, Anh Truong, Eamonn O'Brien, Steven Asquith, Michael Stubbins, Nigel K Spurr, Eric H Lai, and Walt Mahoney. Multiplexed SNP genotyping using the qbead system: a quantum dot-encoded microsphere-based assay. *Nucleic Acids Research*, 31(8):e43, April 2003. ISSN 1362-4962. URL http://www.ncbi.nlm.nih.gov/pubmed/12682378. PMID: 12682378.

[167] Xiao Xu and Stuart K Kim. The early bird catches the worm: new technologies for the caenorhabditis elegans toolkit. *Nature Reviews. Genetics*, 12(11):793–801, 2011. ISSN 1471-0064. doi: 10.1038/nrg3050. URL http://www.ncbi.nlm.nih.gov/pubmed/21969037. PMID: 21969037.

[168] Yoshiyuki Yamada, Takayuki Michikawa, Mitsuhiro Hashimoto, Kazuki Horikawa, Takeharu Nagai, Atsushi Miyawaki, Michael Häusser, and Katsuhiko Mikoshiba. Quantitative comparison of genetically encoded ca indicators in cortical pyramidal cells and cerebellar purkinje cells. *Frontiers in Cellular Neuroscience*, 5:18, 2011. ISSN 1662-5102. doi: 10.3389/fncel.2011.00018. URL http://www.ncbi.nlm.nih.gov/pubmed/21994490. PMID: 21994490.

[169] Ahmet Yildiz, Joseph N Forkey, Sean A McKinney, Taekjip Ha, Yale E Goldman, and Paul R Selvin. Myosin v walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science (New York, N.Y.)*, 300(5628):2061–2065, June 2003. ISSN 1095-9203. doi: 10.1126/science.1084398. URL http://www.ncbi.nlm.nih.gov/pubmed/12791999. PMID: 12791999.

[170] Ahmet Yildiz, Michio Tomishige, Ronald D Vale, and Paul R Selvin. Kinesin walks hand-over-hand. *Science (New York, N.Y.)*, 303(5658):676–678, January 2004. ISSN 1095-9203. doi: 10.1126/science.1093753. URL http://www.ncbi.nlm.nih.gov/pubmed/14684828. PMID: 14684828.

[171] Ofer Yizhar, Lief E Fenno, Thomas J Davidson, Murtaza Mogri, and Karl Deisseroth. Optogenetics in neural systems. *Neuron*, 71(1):9–34, July

2011. ISSN 1097-4199. doi: 10.1016/j.neuron.2011.06.004. URL http://www.ncbi.nlm.nih.gov/pubmed/21745635. PMID: 21745635.

[172] X Yu, B Nguyen, and W O Friesen. Sensory feedback can coordinate the swimming activity of the leech. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 19(11):4634–4643, June 1999. ISSN 1529-2401. URL http://www.ncbi.nlm.nih.gov/pubmed/10341261. PMID: 10341261.

[173] Feng Zhang, Li-Ping Wang, Edward S Boyden, and Karl Deisseroth. Channelrhodopsin-2 and optical control of excitable cells. *Nature Methods*, 3(10):785–792, October 2006. ISSN 1548-7091. doi: 10.1038/nmeth936. URL http://www.ncbi.nlm.nih.gov/pubmed/16990810. PMID: 16990810.

[174] Feng Zhang, Li-Ping Wang, Martin Brauner, Jana F. Liewald, Kenneth Kay, Natalie Watzke, Phillip G. Wood, Ernst Bamberg, Georg Nagel, Alexander Gottschalk, and Karl Deisseroth. Multimodal fast optical interrogation of neural circuitry. *Nature*, 446(7136):633–639, April 2007. ISSN 0028-0836. doi: 10.1038/nature05744. URL http://dx.doi.org/10.1038/nature05744.

[175] Yun Zhang, Hang Lu, and Cornelia I Bargmann. Pathogenic bacteria induce aversive olfactory learning in caenorhabditis elegans. *Nature*, 438 (7065):179–184, November 2005. ISSN 1476-4687. doi: 10.1038/nature04216. URL http://www.ncbi.nlm.nih.gov/pubmed/16281027. PMID: 16281027.

[176] Y Zheng, P J Brockie, J E Mellem, D M Madsen, and A V Maricq. Neuronal control of locomotion in c. elegans is modified by a dominant mutation in the GLR-1 ionotropic glutamate receptor. *Neuron*, 24(2):347–361, October 1999. ISSN 0896-6273. URL http://www.ncbi.nlm.nih.gov/pubmed/10571229. PMID: 10571229.